Université de Montréal

PROBLÈMES D'ESTIMATION DE PARAMÈTRES AVEC RESTRICTION SUR L'ESPACE DES PARAMÈTRES.

par

N'deye Rokhaya Gueye

Département de mathématiques et de statistique
Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de
Philosophiæ Doctor (Ph.D.)
en Statistique

décembre 2003

Université de Montréal
Faculté des études supérieures

Cette thèse intitulée

PROBLÈMES D'ESTIMATION DE PARAMÈTRES AVEC RESTRICTION SUR L'ESPACE DES PARAMÈTRES.

présentée par

N'deye Rokhaya Gueye

a été évaluée par un jury composé des personnes suivantes :

Martin Bilodeau
(président-rapporteur)

François Perron
(directeur de recherche)

Éric Marchand
(co-directeur)

Yves Lepage
(membre du jury)

Glenn Shorrock
(examinateur externe)

(représentant du doyen de la FES)

Thèse acceptée le:

SOMMAIRE

Le présent travail est une contribution à l'étude des problèmes d'estimation de paramètre lorsque l'espace des paramètres est soit une boule de rayon m, $m > 0$ connu ou soit un intervalle $[a, b]$, $a < b$, a et b connus. Nous traitons plusieurs de ces problèmes dans cette thèse.

D'abord, dans le premier chapitre, nous abordons le problème d'estimation du vecteur moyen θ d'une distribution multinormale $N_p(\theta, I_p)$ où θ appartient à une boule de rayon m, $m > 0$ connu, sous la perte quadratique. Nous utilisons les conditions de Marchand et Perron (2001) pour lesquelles, l'estimateur du maximum de vraisemblance est dominé par un estimateur orthogonalement invariant. Ainsi, nous montrons que, pour tout $m \leq c\sqrt{p}$ et pour p assez grand, l'estimateur du maximum de vraisemblance est dominé par l'estimateur de Bayes par rapport à la loi a priori uniforme sur l'espace des paramètres. La démonstration de ce résultat est surtout basée sur la convergence en probabilité de certaines variables aléatoires et l'inégalité d'Amos (1974).

Ensuite, dans le deuxième chapitre, nous considérons le problème d'estimation du paramètre p d'une distribution Binomiale(n, p) lorsque $a \leq p \leq 1-a$, $a < 1/2$. Mais, si nous supposons la longueur de l'intervalle $[a, 1-a]$ égale à $m/\sqrt{n}$ et $p = (1 + \theta/\sqrt{n})/2$, nous nous ramenons alors au problème d'estimation du paramètre θ lorsque $\theta \in [-m, m]$. Ainsi, nous proposons des conditions suffisantes et des conditions nécessaires pour que la loi a priori la moins favorable soit uniforme sur $\{-m, m\}$, sous la perte quadratique et les pertes de la forme $L^*(p, d) = k(p)(d-p)^2$, pour certaines fonctions k déterminées dans la thèse. De plus, nous montrons pour $1 \leq n \leq 4$ que la condition nécessaire et suffisante pour

que la loi a priori la moins favorable soit uniforme sur $\{-m, m\}$, est de la forme $m \leq m_0(n)$.

Enfin, dans le troisième chapitre, nous étudions le problème d'estimation minimax d'un paramètre θ lorsqu'il appartient à un intervalle $[0, m]$, avec m connu. Nous distinguons deux cas, le cas d'une distribution continue de densité de la forme $f(x|\theta) = f_0(x-\theta)I_{[\theta,\infty)}(x)$ et le cas de certains types de distributions discrètes. Dans les deux cas, nous donnons des propriétés générales en utilisant une fonction de perte strictement convexe. Les propriétés générales trouvées dans le cas d'une distribution continue, servirons à traiter les cas particuliers d'une distribution exponentielle sous la fonction de perte LINEX, et d'une distribution uniforme sur $[\theta, \theta+c]$, $c > 0$ donné, sous les fonctions de perte quadratique et LINEX. Les propriétés générales trouvées dans le cas discret serviront à donner des solutions minimax sous la fonction perte quadratique. Nous proposons également des solutions minimax pour le cas d'une distribution de Poisson sous la fonction de perte LINEX.

MOTS CLÉS : admissibilité, distribution Binomiale, distribution exponentielle, distribution de Poisson, distribution Binomiale négative, estimation, estimateur orthogonalement invariant, espace paramétrique restreint, loi a priori la moins favorable, loi a priori uniforme, minimaxité, "Modified Power series", perte quadratique, perte L^q, $q \geq 1$, perte LINEX.

ABSTRACT

This work is a contribution to the problems of estimation when the parameter space is either a ball of radius m, $m > 0$ known or an interval $[a, b]$, $a < b$, a and b known. We treat several of these problems in this thesis.

Initially, in the first chapter, we tackle the problem of estimating the mean vector θ of a multivariate normal $N_p(\theta, I_p)$ distribution where θ belongs to a ball $\Theta(m)$ of radius m, $m > 0$ known, under squared-error loss. We use the conditions of Marchand and Perron (2001) for which the maximum likelihood estimator is dominated by an invariant orthogonally estimator. In particular, we show that, for all $m \leq c\sqrt{p}$ and p large, the maximum likelihood estimator is dominated by the Bayes estimator with respect to a fully uniform prior on the parameter space $\Theta(m)$. The demonstration of this result is mostly especially based on convergence of probability of certain random variables and an inequalities given by Amos (1974).

Then, in the second chapter, we consider the problem of estimating the parameter p of a Binomial(n, p) distribution when $a \leq p \leq 1 - a$, $a < 1/2$. By setting the length of the interval $[a, 1-a]$ equal to $m/\sqrt{n}$, and $p = (1 + \theta/\sqrt{n})/2$, we are reduced then to the problem of estimating parameter θ when $\theta \in [-m, m]$. Thus, we propose sufficient conditions and necessary conditions so that the least favorable prior is uniform on $\{-m, m\}$, under the squared-error loss and loss functions of the form $L(\theta, d) = k(p)(d - p)^2$. Moreover, we show for $1 \leq n \leq 4$ that a necessary and sufficient condition so that the least favorable prior is uniform on $\{-m, m\}$ is of the form $m \leq m_0(n)$.

Lastly, in the third chapter, we study the problem of minimax estimation of a parameter θ when it belongs with an interval $[0, m]$, with m known. We distinguish two cases, the case of a continuous distribution with density of form $f(x|\theta) = f_0(x-\theta)I_{[\theta,\infty)}(x)$, and the case of certain types of discrete distributions. In both cases, we give general properties by using a strict convex loss function. The general properties found in the case of a continuous distribution, will be used to treat the particular cases of an exponential distribution under a LINEX loss function, and of a uniform distribution on $[\theta, \theta + c]$, $c > 0$, under both squared-error and LINEX loss. The general properties found in the discrete case will be used to give minimax solutions for squared-error loss. We also propose minimax solutions in the case of a Poisson distribution under LINEX loss.

KEY WORDS : admissibility, binomial distribution, exponential distribution, Poisson distribution, Geometric distribution, estimation, orthogonally invariant estimator, restricted parametric space, least favorable priors, uniform priors, minimaxity, "Modified Power series" distributions, quadratic loss, L^q loss, LINEX loss.

REMERCIEMENTS

Mes remerciements vont à mon directeur de recherche François Perron et mon co-directeur Éric Marchand pour les problèmes intéressants qu'ils m'ont proposé, pour leur excellent soutien scientifique et financier durant la réalisation de ce travail. À ces grands maîtres, sans qui cette thèse n'aurait pas vu le jour, un grand merci.

Je tiens à exprimer ma reconnaissance aux membres du jury qui ont accepté de lire mon travail.

J'aimerais remercier aussi les professeurs, le personnel administratif et mes collègues étudiants du département de mathématiques et statistique qui ont contribué de près ou de loin à faire progresser mon travail.

Mes remerciements à l'ISM et le département de mathématiques et statistique pour le support financier.

Un très grand merci à mon fils Mame Junior et mon mari SAER dont la présence et le support ont été determinant dans la réussite de ce travail. Encore une fois un très grand merci.

Mes remerciement à ma mère Penda, mes soeurs et frères ainsi que mes amis pour leur soutien moral.

Je tiens à remercier surtout mon père Mame (même s'il n'est plus parmi nous) qui a beaucoup contribué à ma réussite scolaire.

TABLE DES MATIÈRES

LISTE DES FIGURES

LISTE DES TABLEAUX

INTRODUCTION

Les problèmes d'estimation de paramètre avec restriction sur l'espace des paramètres ont suscité un intérêt croissant au cours de la dernière décennie. Plusieurs auteurs se sont intéressés à ces problèmes basés sur des notions classiques d'optimalité à savoir la minimaxité et l'admissibilité.

Notre sujet de recherche porte sur des problèmes d'estimation d'un paramètre θ lorsque ce dernier appartient à une boule de centre 0 et de rayon m, $m > 0$ connu ou à un intervalle $[a, b]$ avec $a < b$, a et b connus. Mais, dans le cas où θ appartient à l'intervalle $[a, b]$, sans perte de généralité, nous prenons $a = 0$ et $b = m$. Notre objectif, dans cette thèse est de construire des estimateurs qui dominent l'estimateur du maximum de vraisemblance et aussi de trouver des estimateurs minimax. Pour rechercher des estimateurs minimax, nous allons procéder de la manière suivante :
- d'abord se limiter aux lois a priori qui concentrent leur masse sur les points de la frontière de l'espace paramétrique,
- ensuite caractériser les estimateurs de Bayes par rapport à ces lois a priori,
- puis trouver pour tout $m > 0$, la loi a priori π_m (si elle existe) pour laquelle la fonction de risque de l'estimateur de Bayes δ_m, associé à π_m, est constante sur la frontière de l'espace paramétrique.
- enfin, utiliser les dérivées première, deuxième et troisième de la fonction de risque de δ_m pour trouver les valeurs de m pour lesquelles la fonction de risque de δ_m atteint son maximum à la frontière de l'espace paramétrique.
Alors, par les résultats de Wald (1950), Brown (1976) et Kempthorne (1987), π_m

est la loi a priori la moins favorable et δ_m est minimax sous certaines conditions.

Cette thèse est divisée en trois chapitres. Le premier chapitre traite le problème d'estimation de la moyenne θ, sous la fonction de perte quadratique, d'une distribution multinormale $N_p(\theta, I_p)$ où θ appartient à l'espace paramétrique $\Theta(m) = \{\theta \in R^p : \|\theta\| \leq m\}$, $m > 0$ fixé. Dans ce chapitre, nous étudions d'abord la minimaxité du problème. Pour ce faire, nous ferons un rappel du résultat de Berry (1990) qui montre, pour tout $p \geq 1$ fixé, l'existence d'un point $m_0(p) > 0$ tel que, pour tout $m \leq m_0(p)$, l'estimateur de Bayes par rapport à la loi a priori uniforme sur la surface de $\Theta(m)$ est minimax. Ensuite, nous traitons l'inadmissibilité de l'estimateur du maximum de vraisemblance. Marchand et Perron (2001) proposent des séries de conditions pour lesquelles l'estimateur du maximum de vraisemblance est dominé. Nous étudions suite aux résultats de ces derniers l'inadmissibilité de l'estimateur du maximum de vraisemblance en traitant le cas où p devient grand. Plus précisément, nous montrons que l'estimateur de Bayes par rapport à la loi a priori uniforme sur tout l'espace paramétrique restreint $\Theta(m)$ domine l'estimateur du maximum de vraisemblance lorsque $m = c\sqrt{p}$ où $c \leq 1$ et p est assez grand.

Le deuxième chapitre porte sur le problème d'estimation d'une proportion p associée à une distribution Binomiale(n, p) lorsque p appartient à un intervalle $[a, 1-a]$ symétrique autour de $1/2$ avec $a < 1/2$. Perron (2003) pose la longueur de l'intervalle $[a, 1-a]$ comme étant égale à $m/\sqrt{n}$. Ceci nous ramène au problème d'estimation du paramètre θ lorsque $\theta \in \Theta(m) = [-m, m]$ en posant $p = (1 + \theta/\sqrt{n})/2$. Nous utilisons certains résultats de Perron (2003) pour proposer des solutions minimax. Plus précisément, nous donnons des conditions suffisantes sur m de la forme $m \leq m_0(n)$, pour tout $n \geq 1$, pour lesquelles la loi a priori uniforme sur $\{a, 1-a\}$ est la moins favorable et l'estimateur de Bayes associé est minimax. Nous utilisons la fonction de perte quadratique et des fonctions de perte de la forme $L^*(p, d) = k(p)(d-p)^2$, pour certaines fonctions k données dans ce deuxième chapitre. Nous proposons aussi des solutions exactes pour $1 \leq n \leq 4$, c'est-à-dire des conditions nécessaires et suffisantes sur m pour lesquelles la loi a priori concentrée sur $\{-m, m\}$ soit la moins favorable.

Dans le troisième chapitre, nous étudions les problèmes d'estimation minimax d'un paramètre θ lorsqu'il appartient à l'intervalle $[0, m]$. Souvent dans la littérature, nous constatons qu'il y a une prédominance de la perte quadratique et donc peu d'étude des pertes asymétriques comme la perte LINEX. De plus, dans le cas de la perte LINEX, les auteurs proposent souvent des conditions seulement suffisantes pour lesquelles la loi a priori la moins favorable soit concentrée sur $\{0, m\}$. Nous avons l'exemple de Wan, Zou et Lee (2000) qui traitent le cas exponentiel pour la perte LINEX.

Dans le troisième chapitre, nos résultats se rapportent à deux classes de distributions qui sont :

(A) : la famille de distributions continues de densité de la forme
$f(x|\theta) = f_0(x-\theta) I_{[\theta,\infty)}(x)$, pour $0 \leq \theta \leq m$,

(B) : la famille de distributions discrètes de fonction de masse p_θ et de support $\{0, 1, 2, ..., s\}$, $s < \infty$ ou $\{0, 1, 2, ...\}$.

Dans les deux familles de distributions, nous utilisons les fonctions de pertes strictement convexes pour développer des propriétés générales pour les estimateurs de Bayes associés à des lois a priori concentrées sur $\{0, m\}$.

Ainsi, pour la famille (A), nous proposons des estimateurs minimax pour une distribution exponentielle avec la perte LINEX et une distribution uniforme sur $[\theta, \theta + c]$, $c \geq m$, avec les pertes quadratique et LINEX. Nos résultats sont analytiques. Plus précisément, nous donnons des conditions nécessaires et suffisantes ou des conditions suffisantes pour lesquelles la loi a priori la moins favorable soit concentrée sur la frontière de l'espace des paramètres.

Pour la famille (B), nous montrons, avec la perte quadratique et sous certaines conditions sur la fonction de masse p_θ, que la loi a priori la moins favorable est concentrée sur $\{0, m\}$ si $m \leq m_0(n)$ où $m_0(n)$ est défini plus loin. Ces résultats viennent des travaux en cours de Marchand et Parsian (2002). Nous appliquons ces résultats à quelques exemples comme la distribution de Poisson, la distribution Binomiale et la distribution Binomiale négative.

En outre, nous traitons aussi la distribution de Poisson avec la perte LINEX. Dans ce cas, dépendemment de la forme de l'assymétrie, nous proposons ou bien des conditions nécessaires et suffisantes ou bien seulement des conditions suffisantes pour que la loi a priori la moins favorable soit concentrée sur $\{0, m\}$.

Dans l'appendice, nous donnons des résultats préliminaires tels des définitions, des propriétés, des théorèmes et des lemmes. Ces résultats préliminaires servent dans les trois chapitres.

Chapitre 1

ESTIMATION DU VECTEUR MOYEN θ D'UNE LOI MULTINORMALE LORSQUE $\|\theta\| \leq m$ AVEC $m > 0$ CONNU

Ce chapitre traite le problème d'estimation de la moyenne θ, sous la fonction de perte quadratique, basé sur l'observation x d'une distribution multinormale $N_p(\theta, I_p)$ où θ appartient à l'espace paramétrique restreint
$\Theta(m) = \{\theta \in I\!R^p : \|\theta\| \leq m\}$, m fixé, $m > 0$.

Plusieurs exemples montrent que l'estimateur naturel $\delta_0(x) = x$ mène à une estimation inadmissible pour le problème non contraint. Par exemple, James et Stein (1961) ont crée une classe dite de type James-Stein donnée par :

$$\Delta_{JS} = \left\{\delta_c(\cdot) : I\!R^p \longrightarrow I\!R^p \quad | \quad \delta_c(x) = \left(1 - \frac{c}{x^t x}\right) x, 0 < c < 2(p-2)\right\}.$$

Ils ont montré, pour $p \geq 3$, que tous les membres de cette classe dominent l'estimateur naturel δ_0. Cependant, l'estimateur sans biais δ_0 ainsi que les estimateurs de type James-Stein ne sont ni minimax, ni admissibles pour le problème contraint.

Mais, la troncature de l'estimateur sans biais δ_0 dans $\Theta(m)$ est donnée par l'estimateur du maximum de vraisemblance $\delta_{evm}(x) = ((m/\|x\|) \wedge 1)x$. Évidemment, cet estimateur domine δ_0 puisque $\|\delta_{evm}(x) - \theta\|^2 \leq \|x - \theta\|^2$, pour tout $x \in I\!R^p$ et pour tout $\theta \in \Theta(m)$ avec inégalité stricte pour tout $x \in \Theta(m)$.

Souvent, les estimateurs du maximum de vraisemblance sont inadmissibles sous la fonction de perte quadratique pour un espace paramétrique restreint. Dans une série d'articles, Charras et van Eeden (1991a, 1991b, 1992, 1994) considèrent le problème d'estimation d'un paramètre θ appartenant à un sous-ensemble fermé et convexe. Ils donnent des conditions pour lesquelles un estimateur "frontière" δ de θ est inadmissible. Ils appliquent souvent leurs résultats aux estimateurs du maximum de vraisemblance dans des espaces paramétriques restreints (c'est le cas de notre problème). Cependant, même s'ils donnent des exemples explicites d'estimateurs qui dominent l'estimateur du maximum de vraisemblance (voir exemple 1.3.1), ces derniers ne sont pas nécessairement de "bons" estimateurs (c'est-à-dire que ces estimateurs ne sont pas monotones) et ils ne se situent pas au coeur de leur recherche. Moors (1981, 1985) considère des problèmes d'estimation invariants dans l'espace des paramètres restreint. Il étudie l'inadmissibilité des estimateurs invariants qui prennent des valeurs sur, ou trop près de la frontière de l'espace des paramètres.

Dans le cas univarié ($p = 1$) et pour $m \leq 1.0567$, Casella et Strawderman (1981) montrent que la loi a priori concentrée sur les valeurs $-m$ et m avec équiprobabilité est la moins favorable et l'estimateur de Bayes associé est non seulement minimax mais domine δ_{evm}. Kempthorne (1987) propose également des solutions minimax pour $p = 1$ et m petit. Bickel (1981) étudie les solutions minimax asymptotiques en m.

Pour le cas multivarié ($p \geq 1$), Berry (1990) montre que l'estimateur de Bayes par rapport à la loi uniforme sur la surface de $\Theta(m)$ est minimax et que cette loi est la moins favorable pour $m \leq m_0(p)$ où $m_0(p)$ est solution d'une équation précisée plus loin (lemme 1.2.2). Pour m petit, Marchand et Perron (2002) proposent des bornes inférieures explicites de $m_0(p)$. Aussi, Marchand et Perron (2001) montrent que si m est petit, tout estimateur de Bayes par rapport à une loi a priori orthogonalement invariante domine δ_{evm}. Lorsque m n'est pas petit, ils proposent des conditions suffisantes pour qu'un estimateur orthogonalement invariant puisse dominer δ_{evm}. En particulier, ils donnent une condition pour laquelle l'estimateur de Bayes δ_U associé à loi a priori π_U uniforme sur $\Theta(m)$ domine δ_{evm}. Nous allons montrer que cette condition se réalise pour $m \leq c\sqrt{p}$ avec $0 \leq c \leq 1$ et pour p assez grand. Ce qui veut dire que pour $m = c\sqrt{p}$ avec $0 \leq c \leq 1$, il existe $M > 0$ tel que pour tout $p > M$, l'estimateur de Bayes par rapport à la loi a priori uniforme sur $\Theta(m)$ domine δ_{evm}.

Dans la première section, nous allons donner les résultats préliminaires, c'est-à-dire des définitions, propriétés, théorèmes et lemmes permettant de démontrer les résultats de ce chapitre. Ensuite, dans la deuxième section, nous étudierons l'estimation minimax pour notre problème. Ces résultats sont de Berry (1990). Puis, nous étudierons l'inadmissibilité de l'estimateur du maximum de vraisemblance dans la troisième section. Nous proposerons des séries de conditions pour lesquelles l'estimateur du maximum de vraisemblance est dominé d'abord par un estimateur orthogonalement invariant et ensuite par un estimateur de Bayes par rapport à une loi a priori orthogonalement invariante. Ces résultats sont de Marchand et Perron (2001). Enfin, dans la quatrième section, nous utiliserons les résultats de ces derniers pour donner des solutions asymptotiques. C'est-à-dire, nous montrons pour $m \leq c\sqrt{p}$ avec $0 \leq c \leq 1$ que l'estimateur du maximum de vraisemblance est dominé asymptotiquement par l'estimateur de Bayes par rapport à la loi a priori uniforme sur $\Theta(m)$.

1.1. Résultats préliminaires

Dans ce chapitre, nous notons $\|X\|$, $\|x\|$ et $\|\theta\|$ par R, r et λ respectivement. Nous considérons les estimateurs orthogonalement équivariants de la forme :

$$\delta_g(x) = \frac{g(r)}{r}x,$$

où la fonction g est appelée multiplicateur de l'estimateur δ_g. Nous définissons la perte encourue en estimant θ par $\delta_g(X)$ comme

$$L(\theta, \delta_g(X)) = \|\theta - \delta(X)\|^2,$$

avec la fonction de risque correspondante égale à

$$R(\theta, \delta_g) = E_\theta[\|\delta_g(X) - \theta\|^2].$$

Soit π_{BU} la loi a priori uniforme sur la surface de $\Theta(m)$. Nous avons (par Perron (1987), Berry (1990) et Robert (1990)) l'estimateur de Bayes δ_{BU} associé à la loi a priori π_{BU} qui est donné par l'expression suivante :

$$\delta_{BU}(x) = \frac{\overline{g}_m(r)}{r} x, \tag{1.1.1}$$

avec $\overline{g}_\lambda(r) = E_\theta[\theta' X/R | R = r] = \lambda \rho_{\frac{p}{2}-1}(\lambda r)$ où $\rho_\nu(t) = I_{\nu+1}(t)/I_\nu(t)$, $\nu > -1$ et I_ν est la fonction de Bessel modifiée.

Nous présentons par le prochain lemme, donné par Watson (1983), les propriétés de $\rho_{\frac{p}{2}-1}$. Ces propriétés nous serons utiles par la suite.

Lemme 1.1.1. *Watson (1983).*

Pour tout $p \geq 1$,

(1) la fonction $\rho_{\frac{p}{2}-1}$ est croissante et concave avec $\rho_{\frac{p}{2}-1}(0) = 0$ et $\lim\limits_{t \to +\infty} \rho_{\frac{p}{2}-1}(t) = 1$,

(2) l'expression $\rho_{\frac{p}{2}-1}(t)/t$ est décroissante en t avec $\lim\limits_{t \to 0} \rho_{\frac{p}{2}-1}(t)/t = 1/p$.

Le prochain lemme, donné par Amos (1974), joue un rôle important dans la démonstration de nos résultats.

Lemme 1.1.2. *Inégalités d'Amos (1974).*

Pour tout $\nu \geq 0$, nous avons :

$$\frac{x}{(\nu+1) + (x^2 + (\nu+1)^2)^{\frac{1}{2}}} \leq \rho_\nu(x) \leq \frac{x}{\nu + (x^2 + (\nu+2)^2)^{\frac{1}{2}}}, \quad \textit{pour tout } x \geq 0.$$

Définition 1.1.1. *Soient δ un estimateur de θ et Y une fonction de X. On appelle risque conditionnel de δ étant donné Y, l'espérance conditionnelle $E_\theta[\|\delta(X) - \theta\|^2 \mid Y(X)]$.*

Les résultats de la section 1.3 sont principalement basés sur les risques conditionnels étant donné $R > m$, $0 < R \leq m$ et $R = r$. Par conséquent, il serait intéressant de chercher la fonction de densité de R. Mais, nous savons que R^2

admet une distribution de khi-deux $\chi_p^2(\lambda^2)$ décentrée. Nous notons par $f_p(\cdot,\lambda)$ la densité de R. Alors pour tout r et $\lambda > 0$, nous avons :

$$f_p(r,\lambda) = r\left(\frac{r}{\lambda}\right)^{\frac{p}{2}-1} I_{\frac{p}{2}-1}(\lambda r)\exp\left(-\frac{\lambda^2+r^2}{2}\right). \tag{1.1.2}$$

Cependant, certains des risques conditionnels ci-dessus mettent en jeu les espérances conditionnelles suivantes :

$$\alpha(m,\lambda) \;=\; E_\theta[\rho_{\frac{p}{2}-1}(\lambda R)|R > m] \quad \text{et} \quad \beta(m,\lambda) = E_\theta\left[\frac{\lambda}{R}\rho_{\frac{p}{2}-1}(\lambda R)|R \leq m\right],$$

aussi bien que les fonctions suivantes :

$$\overline{\alpha}(m) = \sup_{0<\lambda\leq m} \alpha(m,\lambda) \quad \text{et} \quad \overline{\beta}(m) = \sup_{0<\lambda\leq m} \beta(m,\lambda).$$

Des propriétés de $\overline{\alpha}$ et $\overline{\beta}$ sont données dans le prochain lemme.

Lemme 1.1.3. *Marchand et Perron (2001).*

(1) La fonction $\alpha(\cdot,\cdot)$ est croissante en ses deux arguments, par conséquent $\overline{\alpha}(m) = \alpha(m,m)$. De plus, nous avons $\lim_{m\to 0}\overline{\alpha}(m) = 0$ et $\lim_{m\to+\infty}\overline{\alpha}(m) = 1$.

(2) La fonction $\beta(m,\cdot)$ est croissante, par conséquent $\overline{\beta}(m) = \beta(m,m)$. De plus, nous avons $0 \leq \overline{\beta}(m) < m^2/p$ et $\lim_{m\to+\infty}\overline{\beta}(m) \geq 1$.

Finalement, nous allons faire deux rappels, par les deux définitions suivantes, qui nous servirons à la section 1.3 et à la section 1.4 respectivement.

Définition 1.1.2. *Une famille de distributions $\{P_\theta\}$, $\theta \in \Theta$ est dite de famille exponentielle de s paramètres si la densité p_θ de la distribution P_θ par rapport à une certaine mesure μ est de la forme :*

$$p_\theta(x) = \exp\left[\sum_{i=1}^{s}\eta_i(\theta)T_i(x) - B(\theta)\right]h(x),$$

où $\eta_1,\eta_2,...,\eta_s$ et B sont des fonctions réelles des paramètres et pour tout $1 \leq i \leq s$, T_i est une statistique.

Exemple 1.1.1. *Les distributions Normale, Gamma, Khi-deux centrée, Bêta, Binomiale, Poisson, Binomiale Négative appartiennent à des familles exponentielles.*

Définition 1.1.3. *On dit qu'une suite de variables aléatoires X_n converge en probabilité vers X si :*

$$\lim_{n\to\infty} P(|X_n - X| > \epsilon) = 0, \text{ pour tout } \epsilon > 0.$$

1.2. Minimaxité

Il est question, dans cette section, de la recherche d'estimateurs minimax de θ pour notre problème. Pour $p = 1$, il est connu par Wald (1950), Ghosh (1964) et Brown (1976), qu'il existe une unique loi a priori la moins favorable dont le support contient un nombre fini de points et que l'estimateur de Bayes correspondant est minimax. La conjecture est que le nombre de points dans le support de la loi a priori la moins favorable croît avec m. Casella et Strawderman (1981) ont montré que cette loi a priori est concentrée sur les deux points $-m$ et m lorsque $m \leq 1.0567$ et sur les trois points $-m$, 0 et m lorsque m est approximativement compris entre 1.4 et 1.6. Kempthorne (1987) propose des techniques numériques pour évaluer la loi a priori quand on fait varier m. Bickel (1981) fait un changement d'échelle de la loi a priori la moins favorable sur [-1,1] et montre que la limite de la distribution obtenue de ce changement, lorsque $m \to +\infty$, est l'unique solution du problème de minimisation de l'information de Fisher sur toutes les distributions concentrées sur $[-1, 1]$. Il montre que cette limite de distributions G_0 a une densité égale à $g_0(x) = \cos^2(\pi x/2)$, $|x| \leq 1$. Il en déduit que le risque minimax est $\overline{R}(m) = 1 - \pi^2 m^{-2} + o(m^{-2})$, lorsque m est grand. Il faut noter que l'estimateur de Bayes δ_m par rapport à la loi a priori de densité $g_m(x) = g_0(x/m)$, $|x| \leq m$, n'est pas asymptotiquement minimax (car $\limsup_{m\to\infty} R(m, \delta_m) > 1$). Cependant, il construit des estimateurs minimax à partir de G_0. Il généralise ces résultats asymptotiques pour le cas multivarié ($p > 1$).

DasGupta (1985) considère le problème général d'un vecteur de paramètres lorsque l'espace paramétrique restreint Θ est convexe pour une famille générale de distributions. Il montre que, sous des conditions assez générales et pour la perte quadratique, la loi a priori la moins favorable se concentre sur la surface de Θ, lorsque Θ est assez petit. Il traite l'exemple d'une distribution normale $N(\theta, 1)$, lorsque $\Theta = [-m, m]$, $m > 0$.

Berry (1990) propose des conditions pour lesquelles l'estimateur de Bayes par rapport à la loi a priori uniforme sur la surface de $\Theta(m)$ est minimax dans le cas multivarié. Nous allons exposer, dans cette section, les résultats de Berry (1990). Ils sont basés principalement sur deux lemmes et un théorème.

Nous rappelons que l'estimateur de Bayes par rapport à la loi a priori π_{BU} uniforme sur la surface de $\Theta(m)$ est donné par la relation (1.1.1). Par conséquent, la fonction de risque de δ_{BU}, obtenue en utilisant l'estimateur sans biais du risque de Stein, est donnée par :

$$R(\theta, \delta_{BU}) = 2m^2 + \|\theta\|^2 - E_\theta[2mR\rho_{\frac{p}{2}-1}(mR) + m^2\rho^2_{\frac{p}{2}-1}(mR)].$$

Il faut noter que cette fonction est constante lorsque $\|\theta\|$ est fixe. Donc, sans perte de généralité, la fonction de risque sera évaluée en $\theta = (\vartheta, 0, ..., 0)$. On écrira $R(\vartheta, \delta)$, $\vartheta \in \mathbb{R}$ pour représenter $R(\theta, \delta_{BU})$ avec $\theta = (\vartheta, 0, ..., 0)$.

Lemme 1.2.1. *Si X est de loi $N_p(\theta, I_p)$ avec $\|\theta\| \leq m$, alors*

$$\sup_{-m \leq \vartheta \leq m} R(\vartheta, \delta_{BU}) = \max(R(0, \delta_{BU}), R(m, \delta_{BU})).$$

C'est-à-dire que le maximum de la fonction de risque de l'estimateur δ_{BU} est atteint, soit à l'origine soit sur la frontière.

DÉMONSTRATION. En faisant une intégration par partie et un changement de variables, la dérivée de la fonction de risque de δ_{BU} est donnée par :

$$\begin{aligned} \frac{d}{d\vartheta} R(\vartheta, \delta_{BU}) &= 2\vartheta E_\vartheta \left([1 - m^2\rho'_{\frac{p}{2}-1}(mR_1)] - \left[\frac{m}{R_1}\rho_{\frac{p}{2}-1}(mR_1)[m^2\rho'_{\frac{p}{2}-1}(mR_1) + 1] \right] \right), \\ &= 2\vartheta E_\vartheta [A_{p,m}(R_1) - B_{p,m}(R_1)], \end{aligned}$$

où $A_{p,m}(r) = 1 - m^2\rho'_{\frac{p}{2}-1}(mr)$, $B_{p,m}(r) = m\rho_{\frac{p}{2}-1}(mr)[m^2\rho'_{\frac{p}{2}-1}(mr) + 1]/r$ et $R_1^2 \sim \chi^2_{p+2}(\vartheta^2)$. En vertu de la partie (1) du lemme 1.1.1, la fonction $A_{p,m}$ est croissante sur $\mathbb{R}^+$. En appliquant à nouveau le lemme 1.1.1, les fonctions $\rho'_{\frac{p}{2}-1}(mr) + 1$ et $\rho_{\frac{p}{2}-1}(mr)/r$ sont positives et décroissantes sur $\mathbb{R}^+$. Ceci nous montre que $B_{p,m}$ est une fonction décroissante et positive sur $\mathbb{R}^+$. La dérivée de la fonction de risque de δ_{BU} est alors l'intégrale par rapport à r d'une fonction croissante en r. Il existe alors au plus un changement de signe pour l'intégrant, il est soit positif ou passe du négatif au positif. En utilisant alors le résultat de Karlin (1957) «technique de changement de signe», la dérivée par rapport à θ de la fonction de risque de δ_{BU} est soit positive ou passse du négatif au positif lorsque $\theta^2 \geq 0$. La fonction de risque de δ_{BU} est alors soit croissante ou soit décroissante et ensuite croissante en ϑ^2. Ce qui achève la démonstration. □

Lemme 1.2.2. *Pour tout $m \leq m_0(p)$ où m_0 est solution de l'équation $D(m) = 0$ en m avec $D(m)$ donné dans la démonstration, nous avons :*

$$R(m, \delta_{BU}) \geq R(0, \delta_{BU}), \quad \textit{pour tout } m \leq m_0(p).$$

Ce qui veut dire que :

$$R(m, \delta_{BU}) \geq R(\theta, \delta_{BU}), \quad \textit{pour tout } \theta \textit{ tel que } \|\theta\| = m \leq m_0(p).$$

Démonstration. La différence de la fonction de risque de δ_{BU} évaluée en $\theta = 0$ et en $\theta = m$ est donnée par :

$$\begin{aligned} D(m) &= 2m^2 + E_0[[m\rho_{\frac{p}{2}-1}(mR) + R]^2] - E_{m^2}[[m\rho_{\frac{p}{2}-1}(mR) + R]^2], \\ &= 2m^2 + E_{m^2}\left[[m\rho_{\frac{p}{2}-1}(mR) + R]^2 \left(\frac{\exp(m^2/2)m^{(p-2)/2}}{2^{(p-2)/2}\Gamma(p/2)I_{(p-2)/2}(mR)} - 1\right)\right]. \end{aligned}$$

L'intégrant dans l'expression de $D(m)$ est égal à zéro si $\exp(m^2/2)m^{(p-2)/2} = 2^{(p-2)/2}\Gamma(p/2)I_{(p-2)/2}(mr)$ et il fait au moins un changement de signe, soit du positif au négatif, lorsque r^2 varie sur $[0, +\infty)$. En utilisant alors la «technique de changement de signe», $D(m)$ fait au plus un changement de signe du positif au négatif. Ceci nous donne le résultat. □

Ainsi, nous allons utiliser les deux lemmes précédents pour dégager la condition pour laquelle la loi a priori π_{BU} est la moins favorable et l'estimateur δ_{BU} minimax. Ce résultat est donné par le théorème suivant :

Théorème 1.2.1. *Si $m \leq m_0(p)$ où $m_0(p)$ est donné par le lemme 1.2.2 alors π_{BU} est la loi a priori la moins favorable et δ_{BU} est minimax.*

DÉMONSTRATION. Nous savons que, pour $m \leq m_0(p)$, $R(\theta, \delta_{BU})$ varie en fonction de $||\theta||$. Alors nous avons $r(\pi_{BU}, \delta_{BU}) = R(m, \delta_{BU})$. En vertu des deux lemmes précédents, nous avons $r(\pi_{BU}, \delta_{BU}) = \sup_{\theta \in \Theta(m)} R(\theta, \delta_{BU})$. En utilisant alors le théorème 3.3.8 (appendice), nous avons le résultat. □

Des évaluations numériques du $m_0(p)$ sont faites pour $p = 1$ par Casella et Strawderman (1981), pour $p = 2, 3$ par Berry (1990) et pour $4 \leq p \leq 50$ par Perron et Marchand (2002). Ces derniers montrent, pour $p \geq 1$, que $m_0(p) \geq \sqrt{p}$ et que $\lim_{p \to \infty} m_0(p)/\sqrt{p} \approx 1.150963$. Il faut noter que les résultats de Berry (1990) s'appliquent dans le problème d'une distribution multinormale $N_p(\theta, \Sigma)$ avec $\theta' \Sigma^{-1} \theta \leq m$ avec Σ connu.

Robert, Bock et Casella (1990) montrent que si on suppose le problème invariant par transformations orthogonales alors les distributions les moins favorables sont uniformes sur une suite de sphères emboîtées.

1.3. INADMISSIBILITÉ DE L'ESTIMATEUR DU MAXIMUM DE VRAISEMBLANCE

Dans cette section, nous allons étudier l'inadmissibilité de l'estimateur du maximum de vraisemblance δ_{evm} pour notre problème. En quelque sorte, nous proposons des estimateurs qui dominent l'estimateur du maximum de vraisemblance sous certaines conditions. Ces résultats sont de Marchand et Perron (2001). Ils donnent les conditions pour lesquelles δ_{evm} est dominé. Plus précisément, ils montrent que :

(1) si m est assez petit, une large classe d'estimateurs domine δ_{evm}. En particulier, tout estimateur de Bayes par rapport à une loi a priori orthogonalement invariante domine δ_{evm},

(2) si m n'est pas petit, nous avons des conditions générales suffisantes pour qu'un estimateur orthogonalement invariant puisse dominer δ_{evm},

(3) si $m \leq \sqrt{p}$, l'estimateur de Bayes par rapport à la loi a priori uniforme sur la surface de $\Theta(m)$ domine δ_{evm}.

Si nous considérons l'estimateur de Bayes par rapport à une loi a priori orthogonalement invariante, nous avons des conditions sur (m, p) pour que cet estimateur domine δ_{evm}. Nous allons alors se restreindre, dans la première sous-section, aux estimateurs orthogonalement invariants pour établir des conditions pour lesquelles ces derniers dominent δ_{evm}. Ensuite, ces résultats seront appliqués, dans la deuxième sous-section, au cas particulier des estimateurs de Bayes par rapport à une loi a priori orthogonalement invariante. Enfin, sans perte de généralité, pour les lois a priori continues, ces derniers résultats seront appliqués au cas où la loi a priori admet une densité de la forme $Kexp(-h(\|\theta\|^2))$, avec K la constante de normalisation. Les exemples d'une loi a priori uniforme sur $\Theta(m)$, d'une loi a priori uniforme sur la surface de $\Theta(m)$ ainsi que d'autres lois a priori précisées plus tard sont donnés.

1.3.1. Dominance des estimateurs orthogonalement invariants

Les résultats de cette section ont été obtenus en évaluant la différence des risques conditionnels de δ_{evm} et δ_g étant donné ($R \leq m$, $R > m$ et $R = r$).

Remarque 1.3.1. *Soient les deux estimateurs δ_1 et δ_2 et une partition $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_k$ de l'espace échantillonnal. Si $E_\theta[(\delta_1(X) - \theta)^2 - (\delta_2(X) - \theta)^2 | \mathcal{A}_i] \geq 0$ pour tout $\theta \in \Theta(m)$ et pour $i = 1, 2, ..., k$ alors nous avons $R(\theta, \delta_1) - R(\theta, \delta_2) \geq 0$, pour tout $\theta \in \Theta(m)$.*

Ainsi, nous allons chercher des conditions pour lesquelles la différence des risques conditionnels de δ_{evm} et δ_g étant donné $R > m$, $R \leq m$ et $R = r$ est positive. Les trois premiers théorèmes nous donnent ces conditions.

Théorème 1.3.1. *Si g est une fonction croissante sur (m, ∞) et $(2\overline{\alpha}(m) - 1)m < g(r) < m$, pour tout $r \in (m, \infty)$ alors nous avons*

$$E_\theta[L(\theta, \delta_{evm}(X)) - L(\theta, \delta_g(X)) | R > m] > 0, \quad \textit{pour tout } \theta \in \Theta(m).$$

DÉMONSTRATION. Par le lemme 1.1.3, nous avons $(2\overline{\alpha}(m) - 1)m < m$, pour tout $m > 0$. En utilisant le fait que $E_\theta[\theta' X/R | R = r] = \overline{g}_\lambda(r)$, la différence des risques conditionnels de δ_{evm} et δ_g étant donné $R > m$ est donnée par l'expression suivante :

$$E_\theta[L(\theta,\delta_{evm}(X)) - L(\theta,\delta_g(X))|R>m] = E_\theta\left[(m-g(R))\left(g(R) - \left(\left\{2\lambda\rho_{\frac{p}{2}-1}(\lambda R)\right.\right.\right.\right.$$
$$\left.\left.\left.\left. -m\right\}\right)\right)|R>m\right],$$
$$\geq E_\theta[(m-g(R))(g(R)-(2\lambda\alpha(m,R)$$
$$-m))|R>m],$$
$$\geq E_\theta[(m-g(R))(g(R)-(2\overline{\alpha}(m)-1)m)|R>m],$$
$$> 0.$$

La première inégalité vient du fait que les deux fonctions g et $\rho_{p/2-1}$ sont croissantes, donc $\text{Cov}(g(R),\rho_{p/2-1}(\lambda R)|R>m)$ est positif. La deuxième inégalité vient du lemme 1.1.3. □

Exemple 1.3.1. *Charras et van Eeden (1991a) présentent, pour $p=1$, la classe des estimateurs dits de "rétrécissement" de la forme :*

$$\delta_\epsilon(x) = g_\epsilon(r)\frac{x}{r} = xI_{[-m,m]}(x) + (m-\epsilon)I_{(m,\infty)}(x) - (m-\epsilon)I_{(-\infty,-m)}(x),$$

avec $g_\epsilon(r) = (m-\epsilon)I_{(m,\infty)}(r) + rI_{[0,m]}(r)$.
Ces estimateurs sont orthogonalement invariants. Pour ϵ suffisamment petit, nous avons $(2\overline{\alpha}(m)-1)m < (m-\epsilon) < m$ d'où $(2\overline{\alpha}(m)-1)m < g(r) < m$, pour tout $r>m$. D'où, pour ϵ suffisamment petit et pour la fonction de perte quadratique, nous avons $E_\theta[L(\theta,\delta_{evm}(X)) - L(\theta,\delta_g(X))|R>m] > 0$, pour tout $\theta \in \Theta(m)$.

Théorème 1.3.2. *Supposons que $g(r)$ et $r(r-g(r))$ sont des fonctions croissantes en r, pour $r\in(0,m]$. Si $\overline{\beta}(m)<1$ et $(2\overline{\beta}(m)-1)r < g(r) < r$, pour tout $r\in(0,m]$ alors nous avons :*

$$E_\theta[L(\theta,\delta_{evm}) - L(\theta,\delta_g)|R\leq m] > 0, \quad \textit{pour tout } \theta\in\Theta(m).$$

La démonstration de ce théorème est similaire à celle du théorème précédent.

Le corollaire suivant, qui est une conséquence des deux théorèmes précédents, donne la première série de conditions sur g pour lesquelles δ_g domine δ_{evm}.

Corollaire 1.3.1. *Supposons que g est croissant et $g(r)/r$ est décroissant sur $(0,\infty]$. Si $0\leq g(r)\leq r\wedge m$, pour tout $r>0$ et si $g(m)/m \geq 2(\overline{\alpha}(m)\vee\overline{\beta}(m))-1$, alors δ_g domine δ_{evm}.*

Théorème 1.3.3. *Si pour tout $r > r_1$, avec $r_1 = \inf\{r : \overline{g}_m(r)/r < 1, r > 0\}$, $2\overline{g}_m(r) - r \wedge m < g(r) < r \wedge m$ alors nous avons :*

$$E_\theta(L(\theta, \delta_{evm}) - L(\theta, \delta_g)|R = r) > 0, \qquad \textit{pour tout } \theta \in \Theta(m).$$

DÉMONSTRATION. Par le lemme 1.1.1, nous avons $\overline{g}_m(r) \leq m$, donc $r_1 < m$. Il s'en suit que $2\overline{g}_m(r) - r \wedge m < r \wedge m$ si et seulement si $r > r_1$. De plus, nous avons

$$\begin{aligned} E_\theta[L(\theta, \delta_{evm}) - L(\theta, \delta_g)|R = r], &\geq [(r \wedge m - g(r)][g(r) - (2\lambda\rho_{p/2-1}(\lambda r) - r \wedge m)], \\ &\geq [(r \wedge m - g(r)][g(r) - (2\overline{g}_m(r) - r \wedge m)], \\ &> 0. \end{aligned}$$

□

Exemple 1.3.2. *Par le lemme 1.1.1, nous avons $\overline{g}_m(r)/r \leq m^2/p$. Alors, la condition $m \leq \sqrt{p}$, entraîne $\overline{g}_m(r)/r \leq 1$ et $r_1 = 0$. En prenant $g = \overline{g}_m$, la condition du théorème précédente est vérifiée pour tout $r \geq 0$. D'où, l'estimateur de Bayes δ_{BU} domine l'estimateur du maximum de vraisemblance δ_{evm} si $m \leq \sqrt{p}$.* Il faut noter que pour $m \leq \sqrt{p}$, l'estimateur de Bayes δ_{BU} est minimax (par la section 1.2) et domine δ_{evm}.

Le corollaire suivant donne la deuxième série de conditions basées sur (m, p) et le multiplicateur g d'un estimateur orthogonalement invariant δ_g pour que δ_g domine δ_{evm}.

Corollaire 1.3.2. *Si les conditions suivantes sont satisfaites :*

(1) g est croissant,

(2) $0 \leq g(r) \leq r \wedge m$ pour tout $r > 0$,

(3) $m \leq m_1 \wedge \sqrt{p/2}$, où m_1 vérifie l'équation $\overline{\alpha}(m_1) = 1/2$

alors δ_g domine δ_{evm}.

DÉMONSTRATION. C'est une conséquence du théorème 1.3.1 pour $r \in (m, \infty)$ et du théorème précédent, pour $r \in (0, m]$. □

Nous allons à présent utiliser les résultats obtenus dans cette sous-section pour traiter le cas particulier des estimateurs de Bayes associés à des lois a priori orthogonalement invariantes, dans la prochaine sous-section.

1.3.2. Dominance des estimateurs de Bayes associés à des lois a priori orthogonalement invariantes

Avec le théorème suivant, nous commençons par caractériser l'estimateur de Bayes par rapport à la loi a priori π orthogonalement invariante.

Théorème 1.3.4. *Soient $T = \|\theta\|$ et π une loi a priori orthogonalement invariante avec $\pi(\{T=0\}) < 1$. L'estimateur de Bayes δ_π est donné par*

$$\delta_\pi(x) = (g_\pi(r)/r)x,$$

où

$$g_\pi(r) = \frac{\int_0^m t^{-(p/2-2)} I_{\frac{p}{2}}(rt) \exp\left(-t^2/2\right)\sigma(dt)}{\int_0^m t^{-(p/2-1)} I_{\frac{p}{2}-1}(rt) \exp\left(-t^2/2\right)\sigma(dt)}.$$

De plus, g_π est croissant, $0 \le g_\pi \le \overline{g}_m$ et $g_\pi(r)/r \le 1$ si $m \le \sqrt{p}$.

Cette caractérisation de g_π est rendue possible grâce à :

(1) la décomposition d'une variable aléatoire ayant une densité orthogonalement invariante : $\theta = TU$ avec $T = \|\theta\|$ où T est distribué selon une mesure de probabilité σ et $U = T^{-1}\theta$ est distribué uniformément sur la sphère unité de $\mathbb{R}^p$,

(2) les propriétés de l'espérance conditionnelle.

Ce théorème nous garantit la condition (1) du corollaire 1.3.2. Avec l'utilisation des résultats de ce théorème et la condition (3) du corollaire 1.3.2, nous avons la condition (2) du corollaire 1.3.2.

Corollaire 1.3.3. *Si δ_π est un estimateur de Bayes par rapport à une loi a priori orthogonalement invariante alors δ_π domine δ_{evm} lorsque $m \le m_1 \wedge \sqrt{p/2}$.*

Si la condition $m \le m_1 \wedge \sqrt{p/2}$ est satisfaite alors tout estimateur de Bayes par rapport à une loi orthogonalement invariante domine δ_{evm}. Ce résultat est intéressant car la condition $m \le m_1 \wedge \sqrt{p/2}$ ne nécessite que la connaissance de (m, p).

Nous supposons à présent que la loi a priori π a une densité de la forme $K\exp(-h(\|\theta\|^2))$ où K est la constante de normalisation. Nous établissons, par le prochain corollaire, trois conditions sur h et le multiplicateur g de l'estimateur de Bayes δ_π associé à π, pour que δ_π domine δ_{evm}.

Théorème 1.3.5. *Supposons que la loi a priori a une densité de la forme $K\exp(-h(\|\theta\|^2))$ où K est la constante de normalisation. Nous avons alors les résultats suivants :*

(1) si h est croissant alors $g_\pi(r)/r$ est borné supérieurement par 1,

(2) si h est convexe alors $g_\pi(r)/r$ est décroissant en r.

Corollaire 1.3.4. *Supposons que la loi a priori a une densité de la forme $K\exp(-h(\|\theta\|^2))$, où K est la constante de normalisation, telle que :*

- *h est convexe,*
- *$g_\pi(r)/r$ est borné supérieurement par 1,*
- *$g_\pi(m)/m \geq 2(\overline{\alpha}(m) \vee \overline{\beta}(m)) - 1$,*

alors l'estimateur de Bayes δ_π domine δ_{evm}.

Les résultats obtenus de cette section sont appliqués pour les cas particuliers où la loi a priori est uniforme sur $\Theta(m)$, uniforme sur la surface de $\Theta(m)$ et appartient à la famille exponentielle et ils sont résumés dans le tableau de la page suivante.

TAB. 1.1. Conditions pour lesquelles δ_{evm} est dominé.

Estimateur	Conditions A	Conditions B	Conditions C
δ_g	$\cdot\ g$ croissant, $\cdot\ g(r)/r$ décroissant en r, $\cdot\ 0 \le g(r) \le r \wedge m$, $\cdot\ g(m)/m \ge$ $2(\overline{\alpha}(m) \vee \overline{\beta}(m)) - 1$.	$\cdot g$ croissant, $\cdot\ m \le m_1 \wedge \sqrt{p/2}$, $\cdot\ 0 \le g(r) \le r \wedge m$.	$\cdot\ 2\overline{g}_m(r) - r \wedge m < g(r)$, $\cdot\ g(r) < r \wedge m,\ \forall r > r_1$.
δ_π		$\cdot\ m \le m_1 \wedge \sqrt{p/2}$.	
δ_π où $\pi\cdot$ est de la forme $K \exp(-h(\Vert\theta\Vert^2))$	h convexe, $\cdot\ g(r)/r \le 1$, $\cdot\ g(m)/m \ge$ $2(\overline{\alpha}(m) \vee \overline{\beta}(m)) - 1$.		
δ_π où π est uniforme $\cdot$ sur $\Theta(m)$	$\cdot\ g(m)/m \ge$ $2(\overline{\alpha}(m) \vee \overline{\beta}(m)) - 1$.		
δ_π où π est uniforme sur la surface de $\Theta(m)$	$\cdot\ m \le \sqrt{p}$.		
δ_π où π proportionnelle à $e^{-a/2\Vert\theta\Vert^2}$	$\cdot\ m > \sqrt{p}$, $\cdot\ g(m)/m \ge$ $2(\overline{\alpha}(m) \vee \overline{\beta}(m)) - 1$.		

Dans ce tableau,

- δ_g désigne un estimateur orthogonalement invariant,
- δ_π désigne un estimateur de Bayes par rapport à une loi a priori π orthogonalement invariante,
- m_1 vérifie l'équation $\overline{\alpha}(m_1) = 1/2$,
- $r_1 = \inf\{r : \overline{g}_m(r)/r < 1, \quad r > 0\}$

Ces estimateurs doivent satisfaire les conditions A ou B ou C pour dominer δ_{evm}.

1.4. Dominance de l'estimateur de Bayes associé à la loi a priori uniforme sur l'espace paramétrique restreint, lorsque p est assez grand.

Dans cette section, nous considérons l'estimateur de Bayes δ_U par rapport à la loi a priori π_U uniforme sur l'espace paramétrique restreint $\Theta(m)$. D'abord, les deux premières conditions du corollaire 1.3.4 qui sont satisfaites (on pose $h = 0$). Nous obtenons grâce au corollaire 1.3.4 une condition explicite de dominance applicable pour p grand. Ce qui veut dire que, l'estimateur de Bayes δ_U domine δ_{evm} si $g_{\pi_U}(m)/m \geq 2(\overline{\alpha}(m) \vee \overline{\beta}(m)) - 1$. Cependant, pour $1 \leq p \leq 10$, Marchand et Perron (2001) évaluent numériquement cette condition suffisante, qui est de la forme $m \leq m_u(p)$, pour laquelle l'estimateur δ_U domine δ_{evm}. Dans cette section, nous voulons montrer que cette condition se réalise lorsque $m \leq c\sqrt{p}$ où $0 < c \leq 1$ et p assez grand. Pour ce faire, nous allons étudier les comportements de $g_{\pi_U}(m)/m$, $\overline{\alpha}(m)$ et $\overline{\beta}(m)$ lorsque $m = c\sqrt{p}$ et p est assez grand.

Ainsi, nous considérons la variable aléatoire X_p admettant la distribution de khi-deux décentrée de paramètre de décentralité c^2p à p degrés de libertés. Et nous posons $T = \|\theta\|$. Soit π une loi a priori ayant une densité de la forme $K\exp(-h(\|\theta\|^2))$ où K est la constante de normalisation. Marchand et Perron (2001) proposent une représentation de $g_\pi(r)/r$. Elle est donnée par :

$$\begin{aligned}\frac{g_\pi(r)}{r} &= \frac{\int_0^m f_{p+2}(t,r)\exp(-h(t^2))dt}{\int_0^m f_p(t,r)\exp(-h(t^2))dt}, \\ &= E\left[\frac{L}{r}\rho_{\frac{p}{2}-1}(rL)|L \leq m\right],\end{aligned}$$

où la variable aléatoire L a une densité proportionnelle à $f_p(l,r)\exp(-h(l^2))$ avec f_p donné par la relation (1.1.2).

Alors, pour étudier le comportement de $g_{\pi_U}(m)/m$, $\overline{\alpha}(m)$ et $\overline{\beta}(m)$ lorsque $m = c\sqrt{p}$ et p est grand, nous avons besoin d'étudier les convergences en probabilité des variables aléatoires Y_p/p et Z_p/p où $\mathcal{L}(Y_p) = \mathcal{L}(X_p|X_p > c^2p)$ et $\mathcal{L}(Z_p) = \mathcal{L}(X_p|X_p \leq c^2p)$. Les prochains lemmes, nous montrent que ces variables aléatoires convergent en probabilité vers c^2+1 et c^2 (pour $c \leq 1$) respectivement.

Lemme 1.4.1. *Si nous posons $m = c\sqrt{p}$ alors les variables aléatoires X_p/p et Y_p/p convergent en probabilité vers c^2+1.*

DÉMONSTRATION. Nous pouvons écrire $X_p \stackrel{\mathcal{L}}{=} \sum_{i=1}^{p} H_i$, où les H_i sont des variables indépendantes et identiquement distribuées $\chi_1^2(c^2)$. En utilisant la loi faible des grands nombres, $X_p/p \stackrel{\mathcal{L}}{=} \sum_{i=1}^{p} H_i/p \xrightarrow{P} c^2+1$, lorsque $p \to \infty$. Ce qui est équivalent à :

$$\begin{cases} \mathrm{P}\left(X_p \geq p(c^2+1+\varepsilon)\right) \underset{p\to\infty}{\longrightarrow} 0, \text{ pour tout } \varepsilon > 0, \\ \mathrm{P}\left(X_p \leq p(c^2+1-\varepsilon)\right) \underset{p\to\infty}{\longrightarrow} 0, \text{ pour tout } 0 < \varepsilon < 1. \end{cases} \tag{1.4.1}$$

Nous devons maintenant montrer que :

$$\lim_{p\to\infty} \mathrm{P}\left(\left|\frac{Y_p}{p} - c^2 - 1\right| > \varepsilon\right) = 0, \qquad \text{pour tout } \varepsilon > 0.$$

Par la relation (1.4.1), nous avons :

$$\lim_{p\to\infty} \frac{\mathrm{P}\left(X_p \geq p(c^2+1+\varepsilon)\right)}{\mathrm{P}\left(X_p > pc^2\right)} = 0, \quad \text{pour tout } \varepsilon > 0,$$

et

$$\begin{aligned} \lim_{p\to\infty} \frac{P\left(c^2 p \leq X_p \leq p(c^2+1-\varepsilon)\right)}{P\left(X_p > pc^2\right)} &\leq \lim_{p\to\infty} \frac{P\left(X_p \leq p(c^2+1-\varepsilon)\right)}{P\left(X_p > pc^2\right)}, \\ &= 0, \quad \text{pour tout } 0 < \varepsilon < 1. \end{aligned}$$

Ceci nous donne le résultat. □

Lemme 1.4.2. *Si nous posons $m = c\sqrt{p}$ alors la variable aléatoire Z_p/p convergence en probabilité vers c^2, pour tout $c \leq 1$.*

DÉMONSTRATION. Il suffit de montrer que :

$$\mathrm{P}\left(\left|\frac{Z_p}{p} - c^2\right| \geq \varepsilon\right) \underset{p\to\infty}{\longrightarrow} 0, \quad \text{pour tout } \varepsilon > 0.$$

Cette relation est équivalente à la relation suivante :

$$\frac{\mathrm{P}\left(X_p \leq c^2 p(1-\varepsilon)\right)}{\mathrm{P}\left(X_p \leq c^2 p\right)} \underset{p\to\infty}{\longrightarrow} 0, \quad \text{pour tout } \varepsilon > 0. \tag{1.4.2}$$

Considérons maintenant la décomposition suivante :
$X_p \stackrel{\pounds}{=} U+V$, avec $U \sim \chi^2_{(p-1)}$, $V \sim \chi^2_1(c^2p)$, U et V sont indépendantes. Nous avons alors :

$$\begin{aligned}
\mathrm{P}\left(X_p \leq b\right) = \mathrm{P}\left(U+V \leq b\right) &= \int_0^b \int_0^{b-v} f_U(u) f_V(v) du dv, \\
&= b^2 \int_0^1 \int_0^{1-v} f_U(bu) f_V(bv) du dv,
\end{aligned}$$

avec f_U et f_V, les densités respectives des variables aléatoires U et V.
Nous avons alors, pour tout $\epsilon > 0$,

$$\begin{aligned}
\frac{\mathrm{P}\left(X_p \leq c^2p(1-\varepsilon)\right)}{\mathrm{P}\left(X_p \leq c^2p\right)} &= (1-\varepsilon)^2 \frac{\int_0^1 \int_0^{1-v} f_U(c^2p(1-\varepsilon)u) f_V(c^2p(1-\varepsilon)v) du dv}{\int_0^1 \int_0^{1-v} f_U(c^2pu) f_V(c^2pv) du dv}, \\
&= (1-\varepsilon)^2 \int_0^1 \int_0^{1-v} R_p(u,v) \frac{f_U(c^2pu) f_V(c^2pv)}{\int_0^1 \int_0^{1-v} f_U(c^2pu) f_V(c^2pv) du dv} du dv,
\end{aligned}$$

avec

$$R_p(u,v) = \frac{f_U(c^2p(1-\varepsilon)u) f_V(c^2p(1-\varepsilon)v)}{f_U(c^2pu) f_V(c^2pv)}, \quad \text{pour } 0 \leq u+v \leq 1. \qquad (1.4.3)$$

Par la suite, il suffira de montrer que $R_p(u,v) \leq a(p)$ avec $a(p)$ ne dépendant pas de u et v pour tout (u,v) avec $0 \leq u+v \leq 1$ et $\lim_{p \to \infty} a(p) = 0$.
Mais, nous avons les densités des variables aléatoires U et V qui sont données par :

$$f_U(u) = \frac{u^{((p-1)/2-1)} e^{-u/2}}{2^{(p-1)/2} \Gamma((p-1)/2)} \qquad \text{et} \qquad f_V(v) = \frac{1}{2} v^{-1/2} e^{-(v+c^2p)/2} \left(e^{c\sqrt{pv}} + e^{-c\sqrt{pv}}\right).$$

Alors, l'expression (1.4.3) devient :

$$\begin{aligned}
R_p(u,v) &= (1-\varepsilon)^{\frac{p}{2}-2} e^{pc^2 \frac{\varepsilon}{2}(u+v)} \frac{e^{c^2p\sqrt{(1-\varepsilon)v}} + e^{-c^2p\sqrt{(1-\varepsilon)v}}}{e^{c^2p\sqrt{v}} + e^{-c^2p\sqrt{v}}}, \\
&\leq (1-\varepsilon)^{-2} e^{p(c^2\frac{\varepsilon}{2} + \frac{1}{2}log(1-\varepsilon))} \frac{e^{c^2p\sqrt{(1-\varepsilon)v}} + e^{-c^2p\sqrt{(1-\varepsilon)v}}}{e^{c^2p\sqrt{v}} + e^{-c^2p\sqrt{v}}}.
\end{aligned}$$

Comme $\log(1-\varepsilon) \leq -\varepsilon$ pour $0 < \varepsilon < 1$, nous avons :

$$\begin{aligned}
R_p(u,v) &\leq (1-\varepsilon)^{-2} e^{p(c^2-1)\frac{\varepsilon}{2}} \frac{e^{c^2p\sqrt{(1-\varepsilon)v}} + e^{-c^2p\sqrt{(1-\varepsilon)v}}}{e^{c^2p\sqrt{v}} + e^{-c^2p\sqrt{v}}}, \\
&\leq (1-\varepsilon)^{-2} e^{p(c^2-1)\frac{\varepsilon}{2}} e^{c^2p(\sqrt{1-\varepsilon}-1)\sqrt{v}} \frac{1 + e^{-2c^2p\sqrt{(1-\varepsilon)}\sqrt{v}}}{1 + e^{-2c^2p\sqrt{v}}}, \\
&\leq (1-\varepsilon)^{-2} e^{p(c^2-1)\frac{\varepsilon}{2}} \frac{2}{1 + e^{-2c^2p}} = a(p),
\end{aligned}$$

car $e^{c^2p(\sqrt{1-\varepsilon}-1)\sqrt{v}} \leq 1$ et $(1+e^{-2c^2p\sqrt{1-\varepsilon}\sqrt{v}})/(1+e^{-2c^2p\sqrt{v}}) \leq 2/(1+e^{-2c^2p})$ pour $0 \leq v \leq 1$ et $\varepsilon > 0$. Si $c^2 \leq 1$, alors $a(p)$ tend vers 0 lorsque p tend vers l'infini, d'où le résultat. $\square$

Théorème 1.4.1. *Supposons que la loi a priori π admette une loi uniforme sur l'espace paramétrique et que $m = c\sqrt{p}$. Alors, lorsque $p \to \infty$, nous avons :*

(1) $\overline{\alpha}(m)$ converge vers $c/\sqrt{c^2+1}$,

(2) $\overline{\beta}(m)$ et $\overline{g}_\pi(m)/m$ convergent vers $c^2/(1/2+\sqrt{c^4+1/4})$, si $c \leq 1$.

DÉMONSTRATION. (1) Nous savons que :

$$\begin{aligned}\overline{\alpha}(m) &= E\left[\rho_{\frac{p}{2}-1}(c\sqrt{p}\sqrt{X_p})\middle| X_p > c^2p\right]\\ &= E\left[\rho_{\frac{p}{2}-1}(c\sqrt{p}\sqrt{Y_p})\right].\end{aligned}$$

Nous posons $Y = Y_p/p$. En utilisant l'inégalité d'Amos (1974) et pour $\delta \in [1/p, 1/2]$, nous avons le résultat suivant :

$$\frac{c\sqrt{y}}{\frac{1}{2}+\left(c^2y+\frac{1}{4}\right)^{\frac{1}{2}}} \leq \rho_{\frac{p}{2}-1}(cp\sqrt{y}) \leq \frac{c\sqrt{y}}{\frac{1}{2}-\frac{1}{p}+\left(c^2y+\left(\frac{1}{2}+\frac{1}{p}\right)^2\right)^{\frac{1}{2}}} \leq \frac{c\sqrt{y}}{\frac{1}{2}-\delta+\left(c^2y+\frac{1}{4}\right)^{\frac{1}{2}}}.$$

Nous posons :

$$g_\delta(y) = \frac{c\sqrt{y}}{\frac{1}{2}-\delta+\left(c^2y+\frac{1}{4}\right)^{\frac{1}{2}}}.$$

Pour δ fixé, $\delta \in [0, 1/2]$, la fonction g_δ est continue et bornée. De plus, si $\delta \in [0, 1/2]$ alors $g_0(y) \leq \rho_{\frac{p}{2}-1}(cp\sqrt{y}) \leq g_\delta(y)$, pour tout $p > 1/\delta$. Ainsi $g_\delta(Y)$ converge en probabilité vers $g_\delta(1+c^2)$ lorsque p tend vers l'infini et pour c fixé, $g_\delta(1+c^2)$ tend vers $g_0(1+c^2) = c/\sqrt{1+c^2}$ lorsque δ tend vers 0, d'où le résultat.
(2) Nous savons d'autre part que :

$$\begin{aligned}\overline{\beta}(m) &= E\left[\frac{c\sqrt{p}}{\sqrt{X_p}}\rho_{\frac{p}{2}-1}(c\sqrt{p}\sqrt{X_p})\middle| X_p \leq c^2p\right],\\ &= E\left[\frac{c\sqrt{p}}{\sqrt{Z_p}}\rho_{\frac{p}{2}-1}(c\sqrt{p}\sqrt{Z_p})\right],\end{aligned}$$

et

$$\begin{aligned}\frac{g_\pi(m)}{m} &= E\left[\frac{\sqrt{X_p}}{c\sqrt{p}}\rho_{\frac{p}{2}-1}(c\sqrt{p}\sqrt{X_p})\middle| X_p \leq c^2 p\right], \\ &= E\left[\frac{\sqrt{Z_p}}{c\sqrt{p}}\rho_{\frac{p}{2}-1}(c\sqrt{p}\sqrt{Z_p})\right].\end{aligned}$$

En utilisant l'inégalité d'Amos (1974) pour $\overline{\beta}(m)$ et $g_\pi(m)/m$, nous avons :

$$\frac{c^2}{\frac{1}{2}+\left(c^2\frac{Z_p}{p}+\frac{1}{4}\right)^{\frac{1}{2}}} \leq \frac{c\sqrt{p}}{\sqrt{Z_p}}\rho_{\frac{p}{2}-1}(c\sqrt{p}\sqrt{Z_p}) \leq \frac{c^2}{\frac{1}{2}-\frac{1}{p}+\left(c^2\frac{Z_p}{p}+\left(\frac{1}{2}+\frac{1}{p}\right)^2\right)^{\frac{1}{2}}},$$

et

$$\frac{Z_p/p}{\frac{1}{2}+\left(c^2\frac{Z_p}{p}+\frac{1}{4}\right)^{\frac{1}{2}}} \leq \frac{\sqrt{Z_p}}{c\sqrt{p}}\rho_{\frac{p}{2}-1}(c\sqrt{p}\sqrt{Z_p}) \leq \frac{Z_p/p}{\frac{1}{2}-\frac{1}{p}+\left(c^2\frac{Z_p}{p}+\left(\frac{1}{2}+\frac{1}{p}\right)^2\right)^{\frac{1}{2}}}.$$

Pour les mêmes raisons qu'en (1), nous montrons le résultat. □

Théorème 1.4.2. *Si on pose $m = c\sqrt{p}$ avec $0 < c \leq 1$ alors il existe $M > 0$ tel que pour $p > M$, l'estimateur du maximum de vraisemblance est dominé par l'estimateur de Bayes par rapport à la loi a priori uniforme sur $\Theta(m)$.*

Démonstration. Les conditions (1) et (2) du corollaire 1.3.4 sont immédiates. Il reste à vérifier la condition (3) du corollaire 1.3.4. Pour $m = c\sqrt{p}$, nous notons par $\alpha_0 = \lim_{p\to\infty} \overline{\alpha}(m)$ et $\beta_0 = \lim_{p\to\infty} \overline{\beta}(m) = \lim_{p\to\infty} \overline{g}(m)/m$.

(1) Nous montrons d'abord que $\alpha_0 \vee \beta_0 = \alpha_0$.

Mais, nous avons :

$$\begin{aligned}\alpha_0 \vee \beta_0 = \alpha_0 &\iff \frac{c}{\sqrt{c^2+1}} \geq \frac{c^2}{\frac{1}{2}+\sqrt{c^4+\frac{1}{4}}}, \\ &\iff \left(\frac{1}{\sqrt{c^2+1}}\right)^2 \geq \left(\frac{c}{\frac{1}{2}+\sqrt{c^4+\frac{1}{4}}}\right)^2, \\ &\iff \frac{1}{c^2+1} \geq \frac{c^2}{\frac{1}{2}+c^4+\sqrt{c^4+\frac{1}{4}}}, \\ &\iff \frac{1}{2}+\sqrt{c^4+\frac{1}{4}} \geq c^2.\end{aligned}$$

Cette dernière inégalité est vraie pour tout c.

(2) Nous montrons ensuite que $2\alpha_0 - 1 \leq \beta_0$.

Mais, nous avons les équivalences suivantes :

$$
\begin{aligned}
2\alpha_0 - 1 \leq \beta_0 &\iff 2c\left(\frac{1}{2} + \sqrt{c^4 + \frac{1}{4}}\right) \leq \sqrt{c^2+1}\left\{\left(\frac{1}{2} + \sqrt{c^4 + \frac{1}{4}}\right) + c^2\right\}, \\
&\iff 4c^2\left(\frac{1}{2} + \sqrt{c^4 + \frac{1}{4}}\right)^2 \leq (c^2+1)\left(\frac{1}{2} + \sqrt{c^4 + \frac{1}{4}} + c^2\right)^2, \\
&\iff (2c^4 - c^2 + 1)\sqrt{c^4 + \frac{1}{4}} - 2c^6 + 3c^4 - \frac{c^2}{2} + \frac{1}{2} \geq 0, \\
&\iff (2c^4 - c^2 + 1)\left(\sqrt{c^4 + \frac{1}{4}} - c^2 + \frac{4c^4 + c^2 + 1}{4c^4 - 2c^2 + 2}\right) \geq 0.
\end{aligned}
$$

La dernière expression est positive du fait que $2c^4 - c^2 + 1 > 0$ et $\sqrt{c^4 + 1/4} \geq c^2$ pour tout c. $\square$

Chapitre 2

ESTIMATION D'UNE PROPORTION BINOMIALE p LORSQU'ELLE APPARTIENT À L'INTERVALLE $[a, b]$, a ET b CONNUS.

Ce chapitre porte sur le problème d'estimation du paramètre p basé sur l'observation x d'une distribution Binomiale(n, p) où p appartient à l'intervalle $[a, b] \subset (0, 1)$ avec n, a et b connus. Par exemple, Warner (1965) a développé une méthode de collecte de données visant la confidentialité du répondant et par conséquent la réduction et même l'élimination du biais de réponse. Cette méthode porte sur l'estimation de π et $p = \pi P + (1 - \pi)(1 - P)$ avec P connu, $1/2 < P < 1$ et X de distribution Binomiale(n, p).

Dans le cas où $a = 0$, Marchand et MacGibbon (2000) proposent des estimateurs minimax et minimax linéaires pour la fonction de perte quadratique et la fonction de perte normalisée, lorsque b est assez petit. Funo (1991) montre, pour $n \geq 3$, sous la fonction de perte quadratique, l'inadmissibilité de l'estimateur du maximum de vraisemblance si $a = 0$ et $2/n < b < 1$. Charras et van Eeden (1991a) montre, pour $n \geq 3$, l'admissibilité de l'estimateur du maximum de vraisemblance pour $a = 0$ et $b \leq 2/n$.

Dans le cas où l'intervalle $[a, b]$ est symétrique par rapport à $1/2$ (c'est-à-dire $b = 1 - a$ et $a < 1/2$) et sous la fonction de perte quadratique, Moors (1985) montre que l'estimateur de Bayes dépend de la loi a priori correspondante seulement par un nombre fini de ses moments. Il montre aussi que le nombre maximal de points dans le support de la loi a priori la moins favorable ne doit pas excéder $2t + 1$ où t est le plus grand entier inférieur ou égale à $(n + 3)/4$. Il dresse également la liste de tous les estimateurs de Bayes minimax accompagnés des lois a priori les moins favorables pour $1 \leq n \leq 3$. Il donne aussi des résultats numériques pour $4 \leq n \leq 16$.

Berry (1989) utilise une approche plus directe que Moors (1985) et montre, pour $n = 1, 2, 3, 4$, que la loi a priori la moins favorable est supportée par deux ou trois points selon les valeurs sur les bornes de p.

L'estimateur du maximum de vraisemblance est la troncature de la proportion échantillonnale (x/n) dans l'espace des paramètres. Il est donné par
$\delta_{evm}(x) = (1 + (|2x/n - 1| \wedge (1 - 2a)) \text{ sgn } (2x/n - 1))/2$. Il est connu par Sacks (1963), Moors (1985) et Charras et van Eeden (1991a) que cet estimateur est inadmissible car il prend des valeurs sur la borne de l'espace des paramètres.

Cependant, Marchand et MacGibbon (2000) montrent pour $n = 1, 2$ que l'estimateur de Bayes par rapport à une loi a priori symétrique autour de $1/2$, sous certaines conditions, domine l'estimateur du maximum de vraisemblance.

Perron (2003) pose la longueur de l'intervalle $[a, 1 - a]$ comme étant égale à $m/\sqrt{n}$ avec $m < \sqrt{n}$ et se ramène au problème d'estimation du paramètre θ lorsque $\theta \in \Theta(m) = [-m, m]$ en posant $p = (1 + \theta/\sqrt{n})/2$. Il s'inspire alors de Marchand et Perron (2001) et donne des conditions suffisantes pour lesquelles l'estimateur du maximum de vraisemblance est dominé. Il applique ses résultats aux estimateurs proposés par Moors (1985) et Charras et van Eeden (1991a), ainsi à l'estimateur de Bayes dont la densité de la loi a priori est proportionnelle à $(p(1 - p))^{-1}$.

Dans ce chapitre, nous allons traiter seulement le cas où l'intervalle $[a, b]$ est symétrique autour de $1/2$. Le cas où $a = 0$ sera traité dans le chapitre 3. Nous nous intéressons aux estimateurs minimax pour ce problème. Mais généralement, les estimateurs minimax sont difficiles à trouver et les calculs numériques constituent un problème dur. C'est pourquoi nous allons proposer des conditions suffisantes sur m pour lesquelles l'estimateur de Bayes δ_{BU}, associé à la loi a priori π_{BU} uniforme sur la frontière de $\Theta(m)$ est minimax et π_{BU} est la loi a priori la moins favorable, pour la fonction de perte quadratique et des fonctions de perte de la forme $L^*(p,d) = k(p)(d-p)^2$, pour certaines fonctions k déterminées plus loin. Néanmoins, nous avons trouvé des conditions nécessaires et suffisantes sur m pour que π_{BU} soit la loi a priori la moins favorable, pour $1 \leq n \leq 4$. Ces conditions sont de la forme $m \leq m_0(n)$. Nous utiliserons les mêmes notations que Perron (2003).

Ce chapitre est divisé en trois sections. Dans la première section, nous donnerons des résultats préliminaires, c'est-à-dire des définitions, propriétés, théorèmes et lemmes permettant de démontrer nos résultats. Ensuite, dans la deuxième section, nous étudierons l'inadmissibilité de l'estimateur du maximum de vraisemblance. Nous donnerons dans cette section des séries de conditions suffisantes pour lesquelles l'estimateur du maximum de vraisemblance est dominé. Les résultats de cette section viennent surtout de Perron (2003), Moors (1985) et Charras et van Eeden (1991a). Puis dans la troisième section, nous proposerons des solutions (minimax) exactes pour certaines valeurs de n. Ces derniers résultats ont été obtenus en utilisant une méthode directe et différente de celle de Berry (1989) et de Moors (1985). Enfin nous proposerons, dans la quatrième section, des conditions suffisantes de la forme $m \leq m_1(n)$ pour lesquelles la loi a priori uniforme sur la frontière de $\Theta(m)$ est la moins favorable et l'estimateur de Bayes associé à cette loi a priori est minimax.

2.1. Résultats préliminaires

Soit X une variable aléatoire de loi Binomiale(n, p) où p appartient à l'intervalle symétrique $[a, 1-a] \subset (0,1)$ avec $a < 1/2$ et $n \in I\!N^*$. La fonction de masse de X est donnée par :

$$P(X = x|p) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x \in \{0, 1, 2, ..., n\}.$$

Pour se ramener au problème d'estimation du paramètre θ lorsque ce dernier appartient à l'intervalle $[-m, m]$, nous posons la longueur de l'intervalle comme étant égale à $m/\sqrt{n}$ avec $m < \sqrt{n}$. Nous adoptons les notations suivantes :

$\overline{X} = X/n,$

$S = \text{ sgn } (2\overline{X} - 1),$

$R = \sqrt{n}|2\overline{X} - 1|,$

$p = (1 + \theta/\sqrt{n})/2,$

$\lambda = |\theta|,$

$\Theta(m) = \{\theta : |\theta| \leq m\}.$

Nous considérons les estimateurs symétriques, c'est-à-dire les estimateurs qui vérifient la relation $\delta(n - x) = 1 - \delta(x)$, pour $x = 0, 1, ..., n$. Il est claire que ce problème est invariant sous le groupe de transformations $G = \{e, h\}$ avec $\tilde{h}(x) = n - x$ et $\overline{h}(\theta) = 1 - \theta$. Tout estimateur symétrique δ_g peut être associé à une fonction g appelée multiplicateur de l'estimateur δ_g par la formule suivante :

$$\delta_g(x) = \frac{1}{2}\left(1 + g(r)\frac{ms}{\sqrt{n}}\right). \qquad (2.1.1)$$

Par exemple, le multiplicateur de l'estimateur du maximum de vraisemblance est égal à $g_{evm}(r) = 1 \wedge r/m$. La fonction de perte quadratique est utilisée pour donner la fonction de risque de δ_g suivante :

$$\begin{aligned} R(\theta, \delta_g) &= E_\theta[\|\delta_g(X) - p\|^2], \\ &= \frac{1}{4n}E_\theta[\|g(R)mS - \theta\|^2]. \end{aligned} \qquad (2.1.2)$$

La deuxième égalité de l'expression de la fonction de risque de δ_g est obtenue après simplification. Nous allons introduire la fonction ρ_n. Cette fonction joue un rôle important dans la construction de nos estimateurs et de plus elle possède de belles propriétés (voir lemme suivant 2.1.1). Cette fonction est définie de la manière suivante :

$$\begin{aligned} \rho_n : [0, \sqrt{n}] \times [0, \infty) &\longrightarrow [0, 1] \\ (\lambda, r) &\longrightarrow \frac{(1 + \lambda/\sqrt{n})^{\sqrt{n}r} - (1 - \lambda/\sqrt{n})^{\sqrt{n}r}}{(1 + \lambda/\sqrt{n})^{\sqrt{n}r} + (1 - \lambda/\sqrt{n})^{\sqrt{n}r}}. \end{aligned} \qquad (2.1.3)$$

Lemme 2.1.1. *Perron (2003).*

Pour tout $0 \leq \lambda < \sqrt{n}$ *et pour tout* $r \in \Re_n$ *avec*
$\Re_n = \{\sqrt{n}(2k/n - 1) : n/2 \leq k \leq n\}$, *nous avons* $E_\lambda[S|R = r] = \rho_n(\lambda, r)$. *Ainsi, l'estimateur de Bayes* δ_{BU} *par rapport à la loi a priori* π_{BU} *uniforme sur* $\{a, 1-a\}$ *est donné par :*

$$\delta_{BU}(x) = \frac{1}{2}\left(1 + \rho_n(m, r)\frac{ms}{\sqrt{n}}\right).$$

De plus, nous avons les résultats suivants.

(1) La fonction ρ_n *peut s'écrire de la façon suivante :*

$$\rho_n(\lambda, r) = \tanh\left(\log\left(\frac{1 + \lambda/\sqrt{n}}{1 - \lambda/\sqrt{n}}\right)\frac{\sqrt{n}r}{2}\right),$$

pour tout $0 \leq \lambda < \sqrt{n}$ *et* $r \geq 0$.

(2) Pour $0 < \lambda < \sqrt{n}$, *la fonction* $\rho_n(\lambda, \cdot)$ *est croissante et concave avec* $\rho_n(\lambda, 0) = 0$ *et* $\lim_{r\to\infty} \rho_n(\lambda, r) = 1$.

(3) Pour tout $0 \leq \lambda < \sqrt{n}$ *et pour tout* $r \in \Re_n$, *nous avons*

$$\rho_n(\lambda, r) \leq \begin{cases} \lambda r & \textit{si } n \textit{ est impair,} \\ \lambda r/(1 + \lambda^2/n) & \textit{si } n \textit{ est pair.} \end{cases}$$

(4) Pour $r \geq 0$, *la fonction* $\rho_n(\cdot, r)$ *est croissante sur* $[0, \sqrt{n})$.

(5) Pour $0 \leq \lambda < \sqrt{n}$, *l'expression* $\rho_n(\lambda, r)/r$ *est décroissante en* r, *pour* $r > 0$.

Démonstration. En divisant l'expression (2.1.3) par $(1 - \lambda^2/n)^{\sqrt{n}r/2}$ au numérateur et au dénominateur, nous avons le résultat (1).
Ensuite, en utilisant les propriétés de la fonction tanh et le (1), nous avons les résultats de (2), (4) et (5).
Enfin en utilisant le fait que :

$$\rho_n(\lambda, r) = \begin{cases} \lambda r, \text{ pour } r = 0 \text{ et } 1/\sqrt{n}, & 0 \leq \lambda < \sqrt{n}, \\ \lambda r/(1 + \lambda^2/n), \text{ pour } r = 0 \text{ et } 2/\sqrt{n}, & 0 \leq \lambda < \sqrt{n}, \end{cases}$$

et la concavité de $\rho_n(\lambda, \cdot)$, nous avons le résultat (3). □

Les résultats de la section 2.2 (section suivante) sont principalement basés sur les risques conditionnels étant donné $R > m$, $0 < R \leq m$ et $R = r$(voir définition 1.1.1). Comme ces risques conditionnels utilisent la variable aléatoire R, il serait important d'étudier la fonction de masse de R.

Lemme 2.1.2. *Perron (2003).*

Pour $0 \leq \lambda < \sqrt{n}$, la fonction de masse de R est définie sur $\Re_n$. Pour tout $r \in \Re_n$, elle donnée par :

$$f_n(\lambda, r) = \begin{cases} 2^{-n}\binom{n}{n/2}(1-\lambda^2/n)^{n/2} & si\ r = 0, \\ 2^{-n}\binom{n}{\frac{n+\sqrt{n}r}{2}}(1-\lambda^2/n)^{n/2}\left\{\left(\frac{\sqrt{n}+\lambda}{\sqrt{n}-\lambda}\right)^{\sqrt{n}r/2} + \left(\frac{\sqrt{n}-\lambda}{\sqrt{n}+\lambda}\right)^{\sqrt{n}r/2}\right\} & si\ r > 0, \end{cases}$$

$$= \begin{cases} 2^{-n}\binom{n}{n/2}(1-\lambda^2/n)^{n/2} & si\ r = 0, \\ 2^{-n+1}\binom{n}{\frac{n+\sqrt{n}r}{2}}(1-\lambda^2/n)^{n/2}\cosh\left(\frac{\sqrt{n}r}{2}\log\left(\frac{\sqrt{n}+\lambda}{\sqrt{n}-\lambda}\right)\right) & si\ r > 0. \end{cases}$$

Elle a un rapport de vraisemblance monotone en R (c'est-à-dire pour $\lambda_2 < \lambda_1 < \sqrt{n}$, $f_n(\lambda_1, r)/f_n(\lambda_2, r)$ est croissante en r).

DÉMONSTRATION. Nous posons :

$$d_1 = \frac{\sqrt{n}}{2}log\left[\frac{\sqrt{n}+\lambda_1}{\sqrt{n}-\lambda_1}\right] \text{ et } d_2 = \frac{\sqrt{n}}{2}log\left[\frac{\sqrt{n}+\lambda_2}{\sqrt{n}-\lambda_2}\right].$$

Nous pouvons écrire :

$$\frac{f_n(\lambda_1, r)}{f_n(\lambda_2, r)} = \left(\frac{n-\lambda_1}{n-\lambda_2}\right)^{\frac{n}{2}} \frac{e^{d_1 r} + e^{-d_1 r}}{e^{d_2 r} + e^{-d_2 r}}, \quad \text{pour tout} \quad r \geq 0.$$

Nous avons alors

$$\frac{\partial}{\partial r}\left(\frac{e^{d_1 r} + e^{-d_1 r}}{e^{d_2 r} + e^{-d_2 r}}\right) = \frac{(d_1 - d_2)\{e^{(d_1+d_2)r} - e^{-(d_1+d_2)r}\} + (d_1 + d_2)\{e^{(d_1 - d_2)r} - e^{-(d_1-d_2)t}\}}{(e^{d_2 t} + e^{-d_2 t})^2},$$
$$\geq 0.$$

□

Certains des risques conditionnels ci dessus mettent en jeu les espérances conditionnelles suivantes :

$$\alpha_n(\lambda, m) = E_\theta[\rho_n(\lambda, R)|R > m] \quad \text{et} \quad \beta_n(\lambda, m) = E_\theta\left[\frac{\lambda}{R}\rho_n(\lambda, R)|0 < R \leq m\right],$$

aussi bien que les fonctions suivantes :

$$\overline{\alpha}_n(m) = \sup_{0<\lambda\leq m} \alpha_n(m, \lambda) \quad \text{et} \quad \overline{\beta}_n(m) = \sup_{0<\lambda\leq m} \beta_n(m, \lambda).$$

Il faut noter que la définition de β_n ne s'applique que si $P_\lambda[0 < R \leq m] > 0$. Nous proposons les propriétés de α_n et β_n par le lemme suivant.

Lemme 2.1.3. *Perron (2003).*

(1) *La fonction $\alpha_n(\cdot,\cdot)$ est croissante en ses deux arguments, par conséquent $\overline{\alpha}_n(m) = \alpha_n(m,m)$. De plus, nous avons $\lim_{m\to 0} \overline{\alpha}_n(m) = 0$ et $\lim_{m\to\sqrt{n}} \overline{\alpha}_n(m) = 1$.*

(2) *Si n est impair alors $\beta_n(\cdot, m)$ est croissant sur $(0, \sqrt{2n/(n+1)}]$, par conséquent $\overline{\beta}_n(m) = \beta_n(m,m)$ si $m \leq \sqrt{2n/(n+1)}$ et n impair. De plus, nous avons $0 \leq \overline{\beta}_n(m) < m^2$ et $\lim_{m\to\sqrt{n}} \overline{\beta}_n(m) \geq 1$, pour tout $n \geq 1$.*

Nous proposons le lemme suivant. Il est nécessaire dans la démonstration de nos résultats.

Lemme 2.1.4. *Pour tout $r \geq 0$ et pour tout $0 \leq \lambda < \sqrt{n}$, nous avons les résultats suivants :*

(1) $\frac{\partial}{\partial\lambda} \log f_n(\lambda, r) = (1 - \lambda^2/n)^{-1}[r\rho_n(\lambda, r) - \lambda]$,

(2) $\frac{\partial}{\partial\lambda} \rho_n(\lambda, r) = (1 - \lambda^2/n)^{-1} r[1 - \rho^2{}_n(\lambda, r)]$,

(3) $E_\lambda(R\rho_n(\lambda, R)) = \lambda$,

(4) $E_\lambda(R^2) = \lambda^2 + (1 - \lambda^2/n)$,

(5) $E_\lambda[g(R)(R\rho_n(\lambda, R) - \lambda)] = Cov_\lambda[g(R), R\rho_n(\lambda, R)]$.

DÉMONSTRATION.

(1) Pour $r = 0$, la dérivée de f_n par rapport à λ est donnée par :

$$\frac{\partial}{\partial\lambda} \log f_n(\lambda, 0) = \frac{-\lambda}{1 - \lambda^2/n}.$$

Pour $r > 0$, elle est donnée par :

$$\begin{aligned}
\frac{\partial}{\partial\lambda} \log f_n(\lambda, r) &= \frac{2^{-n}\binom{n}{\frac{n+\sqrt{n}r}{2}}\left(1 - \frac{\lambda^2}{n}\right)^{\frac{n}{2}}}{f_n(\lambda, r)} \left[\frac{-\lambda}{1 - \frac{\lambda^2}{n}} \cosh\left(\frac{\sqrt{n}r}{2} \log\left(\frac{\sqrt{n}+\lambda}{\sqrt{n}-\lambda}\right)\right)\right. \\
&\qquad \left. + \frac{\sqrt{n}r}{2} \frac{2\sqrt{n}}{n - \lambda^2} \sinh\left(\frac{\sqrt{n}r}{2} \log\left(\frac{\sqrt{n}+\lambda}{\sqrt{n}-\lambda}\right)\right)\right] \\
&= \frac{n}{n - \lambda^2}\left[-\lambda + r\tanh\left(\frac{\sqrt{n}r}{2} \log\left(\frac{\sqrt{n}+\lambda}{\sqrt{n}-\lambda}\right)\right)\right].
\end{aligned}$$

(2) En dérivant la fonction $\rho_n(\lambda, r)$ par rapport à λ, nous avons :

$$\begin{aligned}
\frac{\partial}{\partial\lambda}\rho_n(\lambda,r) &= \frac{\sqrt{n}r}{2}\frac{2\sqrt{n}}{n-\lambda^2}\left(1-\tanh^2\left(\frac{\sqrt{n}r}{2}\log\left(\frac{\sqrt{n}+\lambda}{\sqrt{n}-\lambda}\right)\right)\right), \\
&= \frac{nr}{n-\lambda^2}(1-\rho_n^2(\lambda,r)).
\end{aligned}$$

La première égalité vient du fait que $\frac{\partial}{\partial x}\tanh(x) = 1 - \tanh^2(x)$ et la deuxième égalité vient de la partie (1) du lemme 2.1.1.

(3) Nous pouvons écrire :

$$\begin{aligned}
E_\lambda(R\rho_n(\lambda,R)) &= E_\lambda(RE_\lambda(S|R)), \\
&= E_\lambda(RS), \\
&= E_\lambda(\sqrt{n}(2\overline{X}-1)), \\
&= \sqrt{n}(2p-1).
\end{aligned}$$

La deuxième égalité vient d'une propriété de l'espérance conditionnelle ($E[E[X|Y]] = E[X]$).

(4) Nous avons aussi :

$$\begin{aligned}
E_\lambda(R^2) &= E_\lambda(n(2\overline{X}-1)^2), \\
&= \mathrm{Var}(2\sqrt{n}\,\overline{X}) + E_\lambda{}^2(\sqrt{n}(2\overline{X}-1)), \\
&= 4p(1-p) + n(2p-1)^2, \\
&= \left(1-\frac{\lambda^2}{n}\right) + \lambda^2.
\end{aligned}$$

(5) Le résultat (5) est une conséquence de la partie (3).

$\square$

La définition suivante nous sera utile à la section 2.3.

Définition 2.1.1. *On dit que la variable aléatoire Y est de loi hypergéométrique $\mathcal{H}(N,b,r)$ si sa fonction de masse est donnée par :*

$$P(Y=y) = \frac{\binom{b}{y}\binom{N-b}{r-y}}{\binom{N}{r}}, \quad (0 \vee (r-N+b)) \le y \le (r \wedge b). \tag{2.1.4}$$

Par exemple, dans une urne composée de N boules blanches et noires, dont b boules blanches. Si on en tire r boules sans remise alors la variable aléatoire Y peut désigner le nombre de boules blanches parmi les r boules tirées.

2.2. L'inadmissibilité de l'estimateur du maximum de Vraisemblance

Dans cette section, nous recherchons des estimateurs admissibles qui dominent l'estimateur du maximum de vraisemblance pour notre problème. Pour un problème d'estimation d'un paramètre $\theta \in \Theta$ ou Θ est un sous-ensemble fermé et convexe de $\mathbb{R}^k$, Charras et van Eeden (1991a) établissent des conditions pour lesquelles un estimateur ne peut être de Bayes et aussi des conditions pour lesquelles cet estimateur est inadmissible. En quelque sorte, ils proposent, dans de nombreux cas, l'inadmissibilité de certains estimateurs dits "estimateurs-frontières" pour l'espace paramétrique convexe et fermé dont l'estimateur du maximum de vraisemblance est fréquemment inclus. Ils traitent notre problème dans leur exemple 5.2 en proposant un estimateur qui prend les mêmes valeurs que l'estimateur du maximum de vraisemblance sauf sur les points frontières.

Soit un problème d'estimation invariant $(\mathcal{X}, \mathcal{A}, \Theta, \mathcal{D}, \mathcal{B}, L)$ (voir appendice) sous l'action d'un groupe fini $G = \{h_1, ..., h_m\}$ de fonctions $g : \mathcal{X} \longrightarrow \mathcal{X}$ avec $\Theta = \{\theta_1, ..., \theta_k\}$ et $L(\theta, d) = \|s(\theta) - d\|^2$ avec s une certaine fonction de $\mathbb{R}^k$ dans $\mathbb{R}^m$, qu'on veut estimer. Nous avons dans ce qui suit quelques résultats de Moors (1985) qui montre l'inadmissibilité des estimateurs équivariants qui prennent des valeurs sur, ou trop près de la frontière de l'espace des paramètres. Nous notons par $\overline{G}$ (respectivement $\tilde{G}$) le groupe de transformation induit par G sur Θ (respectivement sur $\mathcal{D}$). Pour $x \in \mathcal{X}$, nous définissons la fonction suivante :

$$s_x(\theta) = \begin{cases} \dfrac{\sum_{i=1}^{m} f(x|\overline{h}_i(\theta))\tilde{h}_i(s(\theta))}{\sum_{i=1}^{m} f(x|\overline{h}_i(\theta))} \text{ si } \sum_{i=1}^{m} f(x|\overline{h}_i(\theta)) > 0, \overline{h}_i \in \overline{G}, \tilde{h}_i \in \tilde{G}, i = 1, ..., m \\ s(\theta) \text{ si } \sum_{i=1}^{m} f(x|\overline{h}_i(\theta)) = 0. \end{cases}$$

Nous posons $A_x := C(s_x(\Theta))$ (convexe fermé de $s_x(\Theta)$) avec $s_x(\Theta) = \{s_x(\theta) : \theta \in \Theta\}$ et nous avons un des résultats de Moors (1985) donné par le théorème suivant :

Théorème 2.2.1. *Moors (1985).*

Soit δ un estimateur équivariant de $s(\theta)$. Supposons qu'il existe $\theta_0 \in \Theta$ tel que $P_{\theta_0}\{x : \delta(x) \notin A_x\} > 0$ alors δ est strictement dominé par un estimateur équivariant δ_0 où $\delta_0(x)$ est défini pour chaque $x \in \mathcal{X}$ comme la projection de $\delta(x)$ sur A_x.

Nous pouvons dire que tout estimateur équivariant prenant des valeurs en dehors de A_x peut être amélioré en projetant ces valeurs sur A_x.

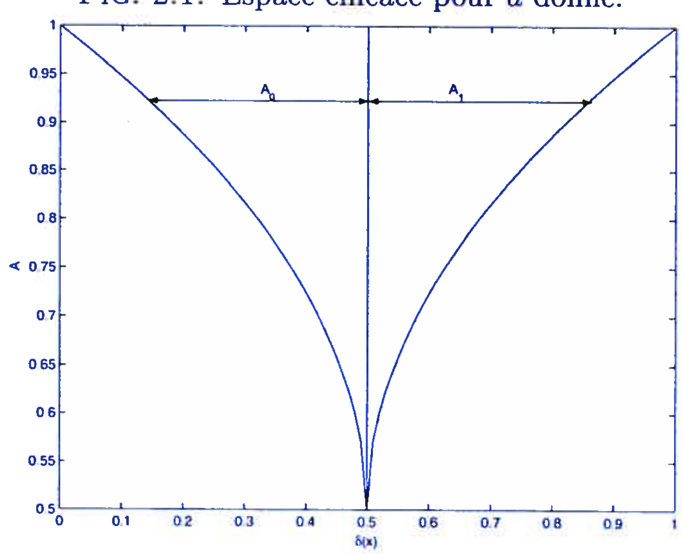
Exemple 2.2.1. *Moors (1985).*

Soient X de distribution Binomiale$(1,p)$,
$L(\theta, d) = (\theta - d)^2$ et $\Theta = (1-a, a)$, avec $1/2 < a \leq 1$. Ce problème est invariant sous le groupe de transformation $G = \{e, h\}$ avec $\tilde{h}(x) = 1 - x$, $\overline{h}(\theta) = 1 - \theta$ et $h^*(d) = 1 - d$. Nous avons :

$$s_x(\theta) = \frac{\theta f(x|\theta) + (1-\theta) f(x|1-\theta)}{f(x|\theta) + f(x|1-\theta)}.$$

Ceci nous donne $s_0(\theta) = 2\theta(1-\theta)$, $s_1(\theta) = \theta^2 + (1-\theta)^2$, $A_0 = [2a(1-a), 1/2]$ et $A_1 = [1/2, a^2 + (1-a)^2]$.

FIG. 2.1. Espace efficace pour a donné.



Nous considérons à présent le problème d'estimation de p d'une distribution Binomiale(n,p) lorsque $p \in [a, 1-a] \subset (0,1)$ avec $a < 1/2$, sous la fonction de perte quadratique. Nous allons d'abord proposer un exemple de Funo (1991) qui démontre l'inadmissibilité de l'estimateur du maximum de vraisemblance pour $a < 1/n$.

Exemple 2.2.2. *Funo (1991).*

Nous supposons que $0 < a < 1/n$. L'estimateur du maximum de vraisemblance est donné par :

$$\delta_{evm}(x) = \begin{cases} a \text{ si } x = 0, \\ x/n \text{ si } x = 1, 2, ..., n-1, \\ 1-a \text{ si } x = n. \end{cases}$$

Pour $0 < \epsilon < \min\{1/n - a, 2a^n(1-2a)/((1-a)^n + a^n), 1-2a\}$, nous définissons l'estimateur suivant :

$$\delta_{\epsilon}(x) = \begin{cases} a + \epsilon \text{ si } x = 0, \\ x/n \text{ si } x = 1, 2, ..., n-1, \\ 1-a-\epsilon \text{ si } x = n. \end{cases}$$

La différence des risques des deux estimateurs δ_{evm} et δ_{ϵ} est donnée par l'expression suivante :

$$\begin{aligned} R(p, \delta_{evm}) - R(p, \delta_{\epsilon}) &= \sum_{x=0}^{n} [(p - \delta_{evm}(x))^2 - (p - \delta_{\epsilon}(x))^2] \binom{n}{x} p^x (1-p)^{n-x}, \\ &= \sum_{x=0}^{n} (\delta_{evm}(x) - \delta_{\epsilon}(x))(\delta_{evm}(x) + \delta_{\epsilon}(x) - 2p) \binom{n}{x} p^x (1-p)^{n-x}, \\ &= \epsilon \zeta(p). \end{aligned}$$

avec $\zeta(p) = (2p - 2a - \epsilon)(1-p)^n - (2p - 2 + 2a + \epsilon)p^n$.

Nous voulons montrer que $\zeta(p) > 0$. Pour cela nous considérons les trois cas suivants.

(1) Supposons que $a \leq p < a + \epsilon/2$. Puisque, $\epsilon \leq 1 - 2a$ alors nous avons $a \leq p < a + \epsilon/2 < 1/2$. Ceci implique les inégalités $0 > 2p - 2a - \epsilon \geq -\epsilon$,

Ce qui nous donne :

$$\begin{aligned}\zeta(p) &= (2p-2a-\epsilon)((1-p)^n - p^n) + 2(1-2a-\epsilon)p^n,\\ &\geq -\epsilon((1-a)^n - a^n) + 2(1-2a-\epsilon)a^n,\\ &> 0.\end{aligned}$$

(2) Supposons que $a+\epsilon/2 \leq p < 1/2$. Dans ce cas, nous avons $2p-2a-\epsilon \geq 0$ et $(1-p)^n \geq p^n$. Il en découle alors les inégalités $\zeta(p) \geq 2(1-2a-\epsilon)p^n > 0$.

(3) Supposons que $1/2 \leq p \leq (1-a)$. La fonction $\zeta(p)$ est symétrique autour de $p = 1/2$ alors par (1) et (2), $\zeta(p) > 0$ pour $1/2 < p \leq 1-a$. De plus, nous avons $\zeta(1/2) = (1-2a-\epsilon)/2^{n-1} > 0$.

Mais, dans les cas particuliers où $n = 1$ et $n = 2$, Marchand et MacGibbon (2000) montrent que l'estimateur de Bayes δ_π par rapport à une loi a priori π symétrique autour de $1/2$ domine l'estimateur du maximum de vraisemblance (δ_{evm}) si et seulement si $\mathrm{Var}_\pi(p) \geq c_n(a)$, $c_1(a) = (1-2a)(1-4a)/4$ et
$c_2(a) = (1-2a)(1-2a-2a^2/(4(1+2a^2-4a^3))$.
Si la loi a priori π est beta(ϵ, ϵ), $\epsilon > 0$, alors δ_π domine δ_{evm} si est seulement si $a \geq a_n$ ou $\epsilon < \epsilon_n(a)$ avec $a_1 = 1/4$, $a_2 = (\sqrt{3}-1)/2$,
$\epsilon_1(a) = (\sqrt{(1-2a)/(1-4a)} - 1)/2$ et
$\epsilon_2(a) = (\sqrt{(1-2a)(1+2a^2-4a^3)/(1-2a-2a^2)} - 1)/2$.

Supposons à présent que la longueur de l'intervalle $[a, 1-a]$ est égale à $m/\sqrt{n}$ (c'est-à-dire $a = (1-m/\sqrt{n})/2$), avec $m < \sqrt{n}$, m donné, Perron (2003) donne des conditions sur (n, m, g) pour lesquelles l'estimateur du maximum de vraisemblance δ_{evm} est dominé par un estimateur équivariant δ_g de la forme (2.1.1). Il montre que si m est assez petit alors tout estimateur δ symétrique autour de $1/2$ et qui rétrécit δ_{evm} autour de $1/2$ le domine. Il propose trois estimateurs qui dominent δ_{evm} pour certaines valeurs de (n, m).
(1) Le premier est tiré de Moors (1985) et son multiplicateur est donné par $g_{mrs}(r) = \rho_n(m, r) \wedge r/m$. Cet estimateur domine δ_{evm}, car il satisfait les conditions du théorème 2.2.3 (donné plus loin). Il coïncide avec l'estimateur de Bayes

par rapport à la loi a priori uniforme sur $\{a, 1-a\}$ si et seulement si $m \leq 1$ pour n impair et $m \leq \sqrt{n/(n-1)}$ pour n pair.

(2) Le second vient de Charras et van Eeden (1991a). Son multiplicateur est égal à $g_{cve}(r) = (m_0 \wedge r)/m$ où

$m_0 = m[(2\overline{\alpha}(m)-1) \vee \sup\{2(\rho_n(m,r) \wedge (r/m)) - r/m) - r/m : 0 \leq r \leq m, r \in \Re_n \cup \{0\}\}]$.

(3) Le troisième est l'estimateur de Bayes par rapport à la loi a priori proportionnel à $(p(1-p))^{-1}$. Son multiplicateur est alors donné par

$g(r) = E[T\rho_n(T,r)/m|R=r]$ où la distribution de T est la distribution conditionnelle de λ étant donné $R=r$.

Ces résultats ont été obtenus d'une part en partitionnant l'espace $\Re_n$ de la manière suivante :

$$\Re_n = \{\{0\} \cap \Re_n\} \cup \{(0,m] \cap \Re_n\} \cup \{(m, \sqrt{n}] \cap \Re_n\}.$$

D'autre part, en évaluant la différence des risques conditionnels de δ_{evm} et δ_g étant donné (soit à $R > m$ ou $R \leq m$ ou $R = r$). On cherche les conditions pour lesquelles cette différence est positive (théorèmes 2.2.2, 2.2.3 ou 2.2.4). Par la remarque 1.3.1, nous en déduisons alors les conditions pour lesquelles δ_g domine δ_{evm}.

Théorème 2.2.2. *Perron (2003).*
Si g est une fonction croissante sur $(m, \sqrt{n}] \cap \Re_n$ et $(2\overline{\alpha}_n(m) - 1) \leq g(r) \leq 1$, pour tout $r \in (m, \sqrt{n}) \cap \Re_n$ alors

$$E_\theta[L(\theta, \delta_{evm}(X)) - L(\theta, \delta_g(X))|R > m] \geq 0, \quad \textit{pour tout } \theta \in \Theta(m).$$

Démonstration. Par le lemme 2.1.3, nous avons $(2\overline{\alpha}_n(m) - 1) < 1$, pour tout $m > 0$. En utilisant la relation (2.1.2) et le fait que $\rho_n(\lambda, r) = E_\lambda[S|R=r]$, la différence des risques conditionnels de δ_{evm} et δ_g étant donné $R > m$ est donnée par l'expression suivante :

$$\begin{aligned} E_\theta[L(\theta,\delta_{evm}(X)) - L(\theta,\delta_g(X))|R > m] &= \frac{m^2}{4n}E_\theta\left[(1-g(R)\left(g(R) - \left(2\frac{\lambda}{m}\rho_n(\lambda,R)-1\right)\right)\right. \\ &\qquad |R>m], \\ &\geq \frac{m^2}{4n}E_\theta\left[(1-g(R))\left(g(R) - \left(2\frac{\lambda}{m}\alpha_n(\lambda,m)-1\right)\right)\right. \\ &\qquad |R>m]. \end{aligned}$$

$$\geq \frac{m^2}{4n} E_\theta[(1-g(R))(g(R)-(2\overline{\alpha}_n(m)-1))|R > m],$$
$$\geq 0.$$

La première inégalité vient du fait que les deux fonctions g et $\rho_n(\lambda, \cdot)$ sont croissantes sur $(m, \sqrt{n}] \cap \Re_n$, donc $\text{Cov}[g(R), \rho_n(\lambda, R)|R > m]$ est positive. La deuxième inégalité vient du lemme 2.1.3. □

De ce théorème, on en tire un corollaire qui donne une première série de conditions pour lesquelles δ_{evm} est dominé par δ_g.

Corollaire 2.2.1. *Perron (2003).*

Supposons que $m < 1/\sqrt{n}$ pour n impair et $m < 2/\sqrt{n}$ pour n pair. Alors, sous les conditions du théorème précédent, nous avons

$$R(\theta, \delta_{evm}) - R(\theta, \delta_g) \geq 0 \text{ pour tout } \theta \in \Theta(m).$$

DÉMONSTRATION. Ce résultat est obtenu du fait que

$$R(\theta, \delta_{evm}) - R(\theta, \delta_g) = E_\theta[L(\theta, \delta_{evm}(X)) - L(\theta, \delta_g(X))|R > m]P_\theta[R > 0],$$

pour tout $\theta \in \Theta(m)$. □

Exemple 2.2.3. *(1) Pour $n = 1$, nous avons $\overline{\alpha}_n(m) = m$. Les conditions du corollaire 2.2.1 sont satisfaites si et seulement si $2m - 1 \leq g(1) \leq 1$. Par exemple, si $g(r) = \rho_n(m, r)$ alors $g(1) = m$. Et lorsque $m \leq 1$, δ_{evm} est dominé par δ_{BU}.*

(2) Pour $n = 2$, $\overline{\alpha}_n(m) = 2\sqrt{2}m/(2+m^2)$ et les conditions du corollaire 2.2.1 sont satisfaites si et seulement si $4\sqrt{2}m/(2 + m^2) - 1 \leq g(\sqrt{2}) \leq 1$. Par exemple, pour $g(r) = \rho_n(m, r)$ nous avons $g(\sqrt{2}) = 2\sqrt{2}m/(2 + m^2)$. Il s'en suit que si $m \leq \sqrt{2}$ alors δ_{evm} est dominé par δ_{BU}.

On se restreint maintenant à l'intervalle $(1/\sqrt{n}, m] \cap \Re_n$.

Théorème 2.2.3. *Perron (2003). Supposons que $m \geq 1/\sqrt{n}$ pour n impair et $m \geq 2/\sqrt{n}$ pour n pair. Soient $g(r)$ et $r(r - g(r))$ des fonctions croissantes en r, pour $r \in (1/\sqrt{n}, m] \cap \Re_n$. Si $\overline{\beta}_n(m) \leq 1$ et $(2\overline{\beta}_n(m) - 1)r/m \leq g(r) \leq r/m$ pour tout $r \in (1/\sqrt{n}, m] \cap \Re_n$ alors nous avons*

$$E_\theta[L(\theta, \delta_{evm}(X)) - L(\theta, \delta_g(X)) \mid 0 < R \leq m] \geq 0, \quad \text{pour tout } \theta \in \Theta(m).$$

DÉMONSTRATION. Puisque $\overline{\beta}_n(m) < m^2$ alors par le lemme 2.1.3, l'ensemble $\{m : \overline{\beta}_n(m) < 1\}$ n'est pas vide. De la même façon que la démonstration du théorème 2.2.2, nous avons le résultat. $\square$

Nous posons $m_- = \sup\{r : r \leq m, r \in \Re_n \cup 0\}$ et $m_+ = \inf\{r : r > m, r \in \Re_n\}$. Nous proposons par le lemme suivant la seconde série de conditions pour lesquelles l'estimateur δ_{evm} est dominé par un estimateur δ_g.

Corollaire 2.2.2. *Perron (2003). Supposons que $0 \leq g(r) \leq 1 \wedge r/m$ pour tout $r \in \Re_n$ avec $g(0) = 0$. Si g est croissant sur $\Re_n$, $g(r)/r$ est décroissant en r pour $r \in \Re_n \setminus \{0\}$, $g(m_-) \geq (2\overline{\beta}_n(m) - 1)m_-/m$ et $g(m_+) \geq 2\overline{\alpha}_n(m) - 1$ alors δ_g domine δ_{evm}.*

DÉMONSTRATION. Puisque g est croissant et borné supérieurement par 1 sur $\Re_n$ alors $g(m_+) \geq 2\overline{\alpha}_n(m) - 1$ entraîne $2\overline{\alpha}_n(m) - 1 \leq g(r) \leq 1$, pour tout $r \in (m, \sqrt{n}] \cap \Re_n$. Nous avons alors les conditions du théorème 2.2.2 qui sont réalisées.

Si $(0, m] \cap \Re_n = \emptyset$ alors la démonstration est complétée.

Sinon, nous savons d'une part que $r(r - mg(r))$ est croissant en $r \in \Re_n$ car $g(r)/r$ décroissant en $r \in \Re_n/\{0\}$. D'autre part $g(m_-) \geq (2\overline{\beta}_n(m) - 1)m_-/m$ implique $(2\overline{\beta}_n(m) - 1)r \leq mg(r) \leq r$, pour tout $0 < r \leq m$, $r \in \Re_n$. Nous vérifions alors les conditions du théorème 2.2.3. $\square$

Théorème 2.2.4. *Perron (2003). Soient $r \in \Re_n$, $\rho_n(m, r) < r/m$ et $2\rho_n(m, r) - (1 \wedge r/m) \leq g(r) \leq 1 \wedge r/m$. Alors, nous avons :*

$$E_\theta[L(\theta, \delta_{evm}(X)) - L(\theta, \delta_g(X)) \mid R = r] \geq 0, \text{ pour tout } \theta \in \Theta(m).$$

DÉMONSTRATION. En utilisant le fait que $\rho_n(\cdot, r)$ est croissant, nous avons :

$$\begin{aligned} 4nE_\theta[L(\theta, \delta_{evm}(X)) - L(\theta, \delta_g(X)) \mid R = r] &= m^2(1 \wedge (r/m) - g(r))(g(r) - (2\lambda\rho_n(\lambda, r) \\ &\quad -1 \wedge (r/m))), \\ &\geq m^2(1 \wedge (r/m) - g(r))(g(r) - (2\rho_n(m, r) \\ &\quad -1 \wedge (r/m))), \\ &\geq 0. \end{aligned}$$

□

Corollaire 2.2.3. *Perron (2003). Supposons que* $0 \leq g(r) \leq 1 \wedge r/m$ *pour tout* $r \in \Re_n$. *Si* g *est croissant sur* $\Re_n$ *et* $m \leq \overline{\alpha}_n^{-1}(1/2) \wedge \sqrt{1/2}$ *alors* δ_g *domine* δ_{evm}.

Corollaire 2.2.4. *Perron (2003). Si* $m_0 \leq m_1 \leq m$ *et* δ_g *est l'estimateur du maximum de vraisemblance pour le nouveau problème où* $|\theta| \leq m_1$ *alors*

$$R(\theta, \delta_{evm}) - R(\theta, \delta_g) \geq 0, \text{ pour tout } \theta \in \Theta(m).$$

Le tableau suivant résume les conditions pour lesquelles δ_g domine δ_{evm}.

TAB. 2.1. Conditions pour lesquelles δ_g domine δ_{evm}.

· $m \leq 1/\sqrt{n}$ pour n impair et $m \leq 2/\sqrt{n}$ pour n pair, · g est croissant sur $(m, \sqrt{n}] \cap \Re_n$, · $(2\overline{\alpha}_n(m) - 1) \leq g(r) \leq 1$, pour tout $r \in (m, \sqrt{n}) \cap \Re_n$.
· $0 \leq g(r) \leq 1 \wedge r/m$, pour tout $r \in \Re_n$ et $g(0) = 0$, · g est croissant sur $\Re_n$, · $g(r)/r$ est décroissant en $r \in \Re_n \setminus \{0\}$, · $g(m_-) \geq (2\overline{\beta}_n(m) - 1)m_-/m$, · $g(m_+) \geq 2\overline{\alpha}_n(m) - 1$.
· $\rho_n(m, r) < r/m$, pour tout $r \in \Re_n$, · $2\rho_n(m, r) - (1 \wedge r/m) \leq g(r), \leq 2 \wedge r/m$, pour tout $r \in \Re_n$.
· $0 \leq g(r) \leq 1 \wedge r/m$ pour tout $r \in \Re_n$, · g est croissant sur $\Re_n$, · $m \leq \overline{\alpha}_n^{-1}(1/2) \wedge \sqrt{1/2}$.
· $m_0 \leq m_1 \leq m$, · δ_g est l'estimateur du maximum de vraisemblance pour le nouveau problème où $\mid \theta \mid \leq m_1$.
· $n = 1$, · $2m - 1 \leq g(1) \leq 1$.
· $n = 2$, · $4\sqrt{2}/(2 + m^2)m - 1 \leq g(\sqrt{2}) \leq 1$.

2.3. Minimaxité : Solutions exactes

Dans cette section, nous nous basons sur le critère de la minimaxité. Par la théorie de l'estimation équivariante (voir Ferguson (1967)), il existe un estimateur non aléatoire (pur) équivariant minimax pour notre problème d'estimation restreint. Notre objectif est de trouver des estimateurs minimax pour notre problème. Souvent, pour résoudre un problème d'estimation minimax, on procède en considérant le problème dual de trouver la loi a priori la moins favorable. La théorie associée à cette dualité s'assure que si les deux problèmes ont des solutions alors l'estimateur de Bayes par rapport à cette loi a priori résout le problème d'estimation minimax. Pour notre problème, il est connu par Wald (1950), Brown (1976) et Kempthorne (1987), qu'il existe une loi a priori la moins favorable, avec un support fini, correspondant à l'estimateur de Bayes minimax. De plus, nous satisfaisons les conditions A à E de l'appendice (conditions de Kempthorne (1987)). Alors par le théorème 3.3.5 et le lemme 3.3.4, si π est la loi a priori la moins favorable qui concentre sa masse sur les points de la frontière alors elle vérifie la relation suivante :

$$\sup_{-m \leq \theta \leq m} R(\theta, \delta_\pi) = R(-m, \delta_\pi) = R(m, \delta_\pi) \tag{2.3.1}$$

Nous commençons par montrer, par le prochain théorème, que le risque d'un estimateur équivariant δ_g est un polynôme en θ^2 de degré $\lfloor \frac{n+1}{2} \rfloor$, où $\lfloor z \rfloor$ est la partie entière de z. En utilisant ce résultat, nous déterminons des conditions pour lesquelles la loi a priori π concentrée sur $\{a, 1-a\}$ avec équiprobabilité est la moins favorable et l'estimateur de Bayes correspond est minimax. Mais, pour démontrer que la fonction de risque de δ_g est un polynôme en θ^2, nous utilisons les résultats du lemme suivant.

Lemme 2.3.1. *Pour tout entier naturel k, nous avons les résultats suivants :*

(1) $(1+\theta)^{n-k}(1-\theta)^k - (1+\theta)^k(1-\theta)^{n-k} = 2\sum_{l=1}^{\lfloor \frac{n+1}{2} \rfloor} \binom{n}{2l-1} G(n,k,2l-1)\theta^{2l-1}$

(2) $(1+\theta)^k(1-\theta)^{n-k} + (1+\theta)^{n-k}(1-\theta)^k = 2\sum_{l=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2l} G(n,k,2l)\theta^{2l}$

où $G(n,k,x)$ est l'espérance de $(-1)^Y$ avec Y de loi hypergéométrique de paramètre (n,k,x) (voir définition 2.1.4).

DÉMONSTRATION. (1) En utilisant la formule du binôme de Newton sur $(1-\theta)^c$ et $(1+\theta)^c$ (avec $c=k$ ou $c=n-k$) , nous avons :

$$\begin{aligned}
(1-\theta)^k(1+\theta)^{n-k}-(1-\theta)^{n-k}(1+\theta)^k &= \sum_{i=0}^{k}\sum_{j=0}^{n-k}\binom{k}{i}\binom{n-k}{j}((-1)^i-(-1)^j)\theta^{i+j},\\
&= \sum_{l=0}^{n}\sum_{i=0}^{k}\binom{k}{i}\binom{n-k}{l-i}((-1)^i-(-1)^{l-i})\theta^l,\\
&= \sum_{l=0}^{n}\left(\sum_{i=0}^{k}-1)^i\frac{\binom{k}{i}\binom{n-k}{l-i}}{\binom{n}{l}}\right)\binom{n}{l}(1-(-1)^l)\theta^l,\\
&= 2\sum_{l=1}^{\lfloor\frac{n+1}{2}\rfloor}G(n,k,2l-1)\binom{n}{2l-1}\theta^{2l-1}.
\end{aligned}$$

La deuxième égalité est obtenue en posant $l=i+j$. La dernière égalité vient du fait que $((-1)^l-1)=0$ si l est pair.

Le résultat (2) s'obtient de façon similaire au résultat (1). □

Théorème 2.3.1. *Sous la perte quadratique, le risque d'un estimateur δ_g peut s'écrire de la façon suivante :*

$$\begin{aligned}
R(\theta,\delta_g) &= \frac{1}{2^{n-1}}\sum_{i=0}^{\lfloor\frac{n}{2}\rfloor+1}A_i(n)\theta^{2i}, \quad \textit{avec } A_0(n)=\sum_{k=0}^{\lfloor\frac{n}{2}\rfloor}\binom{n}{k}\frac{m^2}{n}g^2\left(\frac{n-2k}{\sqrt{n}}\right),\\
A_1(n) &= \frac{m}{\sqrt{n}}\sum_{k=0}^{\lfloor\frac{n}{2}\rfloor}\binom{n}{k}g\left(\frac{n-2k}{\sqrt{n}}\right)\left\{\frac{m}{\sqrt{n}}\binom{n}{2}G(n,k,2)g\left(\frac{n-2k}{\sqrt{n}}\right)\right.\\
&\qquad\left.-2nG(n,k,1)\right\}+2^{n-1},\\
A_l(n) &= \frac{m\sqrt{n}}{n^l}\sum_{k=0}^{\lfloor\frac{n}{2}\rfloor}\binom{n}{k}g\left(\frac{n-2k}{\sqrt{n}}\right)\left\{\frac{m}{\sqrt{n}}\binom{n}{2l}G(n,k,2l)g\left(\frac{n-2k}{\sqrt{n}}\right)\right.\\
&\qquad\left.-2\binom{n}{2l-1}G(n,k,2l-1)\right\}, \textit{pour } l=2,\ldots,\left\lfloor\frac{n}{2}\right\rfloor,\\
A_{\lfloor\frac{n}{2}\rfloor+1}(n) &= \begin{cases}-2m(\sqrt{n})^{-n}\sum_{k=0}^{\lfloor\frac{n}{2}\rfloor}\binom{n}{k}g\left(\frac{n-2k}{\sqrt{n}}\right)G(n,k,n) \textit{ si } n \textit{ est impair },\\ 0 \textit{ si } n \textit{ est pair }.\end{cases}
\end{aligned}$$

DÉMONSTRATION. Nous savons que :

$$
\begin{aligned}
4nR(\theta,\delta_g) &= E_\theta\left\{mSg(R)-\theta\right\}^2, \\
&= \sum_{k=0}^{\lfloor \frac{n}{2}\rfloor}\left\{mSg\left(\frac{n-2k}{\sqrt{n}}\right)-\theta\right\}^2\binom{n}{k}\frac{1}{2^n}\left(1+\frac{\theta}{\sqrt{n}}\right)^k\left(1-\frac{\theta}{\sqrt{n}}\right)^{n-k}
\end{aligned}
$$
$$
+\sum_{k=\lfloor \frac{n}{2}\rfloor+1}^{n}\left\{\frac{mS}{\sqrt{n}}g\left(\frac{2k-n}{\sqrt{n}}\right)-\theta\right\}^2\binom{n}{k}\frac{1}{2^n}\left(1+\frac{\theta}{\sqrt{n}}\right)^k\left(1-\frac{\theta}{\sqrt{n}}\right)^{n-k}.
$$

En utilisant la symétrie de l'estimateur δ_g (c'est-à-dire $\delta_g(n-x)=1-\delta_g(x)$), nous avons le résultat suivant.

$$
4nR(\theta,\delta_g) = \sum_{k=0}^{\lfloor \frac{n}{2}\rfloor}\binom{n}{k}\frac{1}{2^n}\left\{m^2g^2\left(\frac{n-2k}{\sqrt{n}}\right)\left[\left(1+\frac{\theta}{\sqrt{n}}\right)^k\left(1-\frac{\theta}{\sqrt{n}}\right)^{n-k}\right.\right.
$$
$$
\left.+\left(1+\frac{\theta}{\sqrt{n}}\right)^{n-k}\left(1-\frac{\theta}{\sqrt{n}}\right)^{k}\right]+2m\theta g\left(\frac{n-2k}{\sqrt{n}}\right)\left[\left(1+\frac{\theta}{\sqrt{n}}\right)^k\left(1-\frac{\theta}{\sqrt{n}}\right)^{n-k}\right.
$$
$$
\left.\left.-\left(1+\frac{\theta}{\sqrt{n}}\right)^{n-k}\left(1-\frac{\theta}{\sqrt{n}}\right)^{k}\right]\right\}+\theta^2.
$$

En appliquant les résultats du lemme précédent, nous avons :

$$
R(\theta,\delta_g)=\frac{1}{2^{n-1}}\sum_{k=0}^{\lfloor \frac{n}{2}\rfloor}\binom{n}{k}m^2g^2\left(\frac{n-2k}{\sqrt{n}}\right)\sum_{l=0}^{\lfloor \frac{n}{2}\rfloor}\binom{n}{2l}\left(\frac{\theta}{\sqrt{n}}\right)^{2l}G(n,k,2l)
$$
$$
-2m\frac{\sqrt{n}}{2^{n-1}}\sum_{k=0}^{\lfloor \frac{n}{2}\rfloor}\binom{n}{k}g\left(\frac{n-2k}{\sqrt{n}}\right)\sum_{l=1}^{\lfloor \frac{n+1}{2}\rfloor}\binom{n}{2l}\left(\frac{\theta}{\sqrt{n}}\right)^{2l}G(n,k,2l-1)+\theta^2.
$$

□

Il faut noter que $A_0(n)\geq 0$, pour tout $n\geq 1$. Le multiplicateur de l'estimateur de Bayes uniforme sur la frontière de $\theta(m)$ peut s'écrire de la façon suivante :

$$
g_{BU}\left(\frac{n-2k}{\sqrt{n}}\right)=\frac{\left(1+\frac{m}{\sqrt{n}}\right)^{n-2k}-\left(1-\frac{m}{\sqrt{n}}\right)^{n-2k}}{\left(1+\frac{m}{\sqrt{n}}\right)^{n-2k}+\left(1-\frac{m}{\sqrt{n}}\right)^{n-2k}}.
$$

(1) Pour $n=1$, la fonction de risque est linéaire en θ^2 avec $A_1(1)=1-2mg(1)$. En prenant $g=g_{BU}$, nous avons $A_1(1)=1-2m^2$. La fonction de risque en δ_{BU} atteint son maximum en m si et seulement si $m^2\leq 1/2$. Nous avons alors $m_0(1)=1/\sqrt{2}\approx 0.7071$. L'estimateur de

Bayes δ_{BU} est alors minimax et π_{BU} est la moins favorable si et seulement si $m \leq m_0(1)$.

(2) Pour $n = 2$, la fonction de risque est linéaire en θ^2 avec $A_1(2) = m^2 g^2(\sqrt{2})/2 - 4mg(\sqrt{2})/\sqrt{2} + 2$. En prenant $g = g_{BU}$, nous avons $A_1(2) = -2(m^4 + 4m^2 - 4)/(2+m^2)^2$. La fonction de risque en δ_{BU} atteint son maximum en m si et seulement si $m^2 \leq 2(\sqrt{2} - 1)$. L'estimateur de Bayes δ_{BU} est minimax et π_{BU} est la moins favorable si et seulement si $m \leq m_0(2)$ avec $m_0(2) = \sqrt{2(\sqrt{2} - 1)} \approx 0.9102$.

(3) Pour $n = 3$, la fonction de risque est un polynôme de degré deux en θ^2 avec $A_2(3)$ positif. Comme $A_2(3)$ est positif alors la fonction de risque a la forme d'un "U" ou "W". Cette fonction atteint son maximum en m si et seulement si $R(m, \delta_{BU}) \geq R(0, \delta_{BU})$, c'est-à-dire $A_1(3) + m^2 A_2(3) \geq 0$. Mais, la fonction $A_1(3) + m^2 A_2(3)$ passe du positif au négatif. D'où la loi a priori π_{BU} est la moins favorable et l'estimateur de Bayes δ_{BU} est minimax si et seulement si $m \leq m_0(3)$ avec $m_0(3) \approx 0.9215$ est solution de l'équation $A_1(3) + m^2 A_2(3) = 0$.

(4) Pour $n = 4$, similairement au cas $n = 3$, nous montrons que l'estimateur de Bayes δ_{BU} est alors minimax et π_{BU} est la loi a priori la moins favorable si et seulement si $m \leq m_0(4)$ avec $m_0(4) \approx 0.9628$ est solution de l'équation $A_1(4) + m^2 A_2(4) = 0$.

2.4. Minimaxité pour m assez petit

Dans la section précédente, nous avons trouvé des conditions nécessaires et suffisantes, de la forme $m \leq m_0(n)$, pour que la loi a priori π_{BU} soit la moins favorable et l'estimateur de Bayes δ_{BU} correspondant soit minimax pour $1 \leq n \leq 4$. Pour $n \geq 5$, il est très difficile de trouver des conditions nécessaires et suffisantes sur m pour avoir π_{BU} la loi a priori la moins favorable. C'est pourquoi, nous allons proposer des conditions suffisantes sur m pour la fonction de perte quadratique ($L(p,d) = (d-p)^2$) et les fonctions de perte de la forme $L^*(p,d) = (c-\lambda^2)^{-1}L(\theta,d)$ avec $c \geq m^2$ pour notre problème. Pour ce faire, nous allons commencer par trouver des conditions pour lesquelles la fonction de risque d'un estimateur δ_g est croissante en λ, par les théorèmes 2.4.1 et 2.4.4 respectivement aux fonctions de perte ci-dessus. Sous ces conditions, nous avons la fonction de risque de δ_g qui atteint son maximum sur la frontière. Nous en déduisons alors, par les théorèmes 2.4.2 et 2.4.5, les conditions suffisantes pour lesquelles π_{BU} soit la loi a priori la moins favorable pour les deux fonctions L et L^*.

Pour toute fonction g positive et croissante sur $\Re_n$, nous définissons la fonction h par :

$$h(r) = \begin{cases} g(2/\sqrt{n})/(2/\sqrt{n}) & \text{si } r \in \Re_n \text{ et } r = 0, \\ g(r)/r & \text{si } r \in \Re_n \text{ et } r > 0. \end{cases} \tag{2.4.1}$$

Théorème 2.4.1. *Supposons que la fonction h soit décroissante sur $\Re_n$ et que $g(0) = 0$. Nous avons les résultats suivants :*

(1) $4n\frac{\partial}{\partial\lambda}R(\lambda,\delta_g) \geq 2\lambda E_\lambda[1 - 2mh(R)]$, *pour tout* $\lambda \in [0, m \wedge 1]$.

(2) Si $mE_0[h(R)] \leq 1/2$ *alors la fonction de risque* $R(\lambda,\delta_g)$ *est croissante en* λ, *pour* $\lambda \in [0, m \wedge 1]$.

(3) Si $mh(r_0) \leq 1/2$ *avec* $r_0 = \min\{r : r \in \Re_n\}$ *alors la fonction de risque* $R(\lambda,\delta_g)$ *est croissante en* λ, *pour* $\lambda \in [0, m \wedge 1]$.

Démonstration. (1) Nous pouvons écrire :

$$\begin{aligned} 4nR(\lambda,\delta_g) &= E_\theta[m^2g^2(R) + \lambda^2 - 2mS\theta g(R)], \\ &= E_\lambda[m^2g^2(R) + \lambda^2 - 2m\lambda g(R)\rho_n(\lambda,R)]. \end{aligned} \tag{2.4.2}$$

La dérivée première du risque de δ_g est donnée par :

$$
\begin{aligned}
\frac{\partial}{\partial\lambda}R(\lambda,\delta_g) &= \frac{1}{4n}\frac{\partial}{\partial\lambda}E_\lambda\left[m^2g^2(R)+\lambda^2-2m\lambda g(R)\rho_n(\lambda,R)\right],\\
&= \frac{1}{4n}E_\lambda\left[\frac{\partial}{\partial\lambda}\left\{m^2g^2(R)+\lambda^2-2m\lambda g(R)\rho_n(\lambda,R)\right\}\right.\\
&\qquad\left.+\left\{m^2g^2(R)+\lambda^2-2m\lambda g(R)\rho_n(\lambda,R)\right\}\frac{\partial}{\partial\lambda}\log f_n(\lambda,R)\right],\\
&= \frac{1}{4n}E_\lambda\left[2\lambda-2mg(R)\left(\rho_n(\lambda,R)+\lambda\left(1-\frac{\lambda^2}{n}\right)^{-1}R(1-\rho_n^2(\lambda,R))\right)\right.\\
&\qquad\left.+\left(1-\frac{\lambda^2}{n}\right)^{-1}\left(m^2g^2(R)+\lambda^2-2m\lambda g(R)\rho_n(\lambda,R)\right)\left(R\rho_n(\lambda,R)-\lambda\right)\right],\\
&= \frac{1}{4n}\left\{2E_\lambda[\lambda-mg(R)\rho_n(\lambda,R)]+\left(1-\frac{\lambda^2}{n}\right)^{-1}E_\lambda[(m^2g^2(R)+\lambda^2)(R\rho_n(\lambda,R)-\lambda)]\right.\\
&\qquad\left.-2m\lambda\left(1-\frac{\lambda^2}{n}\right)^{-1}E_\lambda[g(R)(R-\lambda\rho_n(\lambda,R))]\right\},\\
&= \frac{1}{4n}\left\{2E_\lambda[\lambda-mg(R)\rho_n(\lambda,R)]+\left(1-\frac{\lambda^2}{n}\right)^{-1}\mathrm{Cov}_\lambda(m^2g^2(R)+\lambda^2,R\rho_n(\lambda,R))\right.\\
&\qquad\left.-2m\lambda\left(1-\frac{\lambda^2}{n}\right)^{-1}E_\lambda[h(R)[R^2-\lambda R\rho_n(\lambda,R)-(1-\lambda^2/n)]]-2m\lambda E_\lambda(h(R))\right\},\\
&= \frac{1}{4n}\left\{2E_\lambda[\lambda-mg(R)\rho_n(\lambda,R)-m\lambda h(R)]+\left(1-\frac{\lambda^2}{n}\right)^{-1}\right.\\
&\qquad\mathrm{Cov}_\lambda(m^2g^2(R)+\lambda^2,R\rho_n(\lambda,R))\\
&\qquad\left.-2m\lambda\left(1-\frac{\lambda^2}{n}\right)^{-1}\mathrm{Cov}_\lambda(h(R),R^2-\lambda R\rho_n(\lambda,R))\right\}.
\end{aligned}
$$

La troisième égalité vient des résultats (1) et (2) du lemme (2.1.4). Les cinquième et sixième égalités viennent des résultats (3), (4) et (5) du lemme (2.1.4).

Nous savons pour $\lambda \geq 0$ que $r\rho_n(\lambda,r)$ est croissante en r (lemme 2.1.1 (1)). Alors, si la fonction g est positive et croissante, $\mathrm{Cov}_\lambda(m^2g^2(R)+\lambda^2,R\rho_n(\lambda,R))$ est positif. Nous pouvons écrire $r^2-\lambda r\rho_n(\lambda,r) = r^2(1-\lambda\rho_n(\lambda,r)/r)$. Comme l'expression $(1-\lambda\rho_n(\lambda,r)/r)$ est croissante en r (lemme 2.1.1 (4)) et positive pour $\lambda \leq 1$ alors l'expression $\mathrm{Cov}_\lambda(h(R),R^2-\lambda R\rho_n(\lambda,R))$ est positive.

Nous avons alors :

$$\begin{aligned}\frac{\partial}{\partial\lambda}R(\lambda,\delta_g) &\geq \frac{1}{4n}2E_\lambda[\lambda - mg(R)\rho_n(\lambda,R) - m\lambda h(R)],\\ &\geq \frac{1}{2n}\lambda E_\lambda[1-2mh(R)] \quad \text{car} \quad E_\lambda(g(R)\rho_n(\lambda,R)) \leq \lambda E_\lambda(h(R)).\end{aligned}$$

(2) Par le résultat (1), nous avons :

$$\frac{\partial}{\partial\lambda}R(\lambda,\delta_{BU}) \geq \frac{1}{2n}\lambda E_\lambda[1-2mh(R)].$$

Comme h est décroissante et que la famille des distributions de R est à rapport de vraisemblance monotone en r, nous avons $E_\lambda[h(R)] \leq E_0[h(R)]$. Il en découle de cette dernière inégalité que si $mE_0[h(R)] \leq 1/2$ alors $\frac{\partial}{\partial\lambda}R(\lambda,\delta_g) \geq 0$, pour tout $\lambda \in [0, m \wedge 1]$.

(3) Comme la fonction $h(r)$ est décroissante sur $\Re_n$ alors $E_0[h(R)] \leq h(r_0)$. En utilisant la partie (2), nous avons le résultat. □

Théorème 2.4.2. *Soit $h(m,r) = \rho_n(m,r)/r$, pour tout $r \in \Re_n$, $r \neq 0$ et $h(m,0) = g(2/\sqrt{n})/(2/\sqrt{n})$ si $r \in \Re_n$ et $r = 0$. Nous avons les résultats suivants :*

(1) Pour $m \leq m_1(n) \wedge 1$ avec $m_1(n)$ solution de l'équation $mE_0[h(m,R)] = 1/2$, l'estimateur de Bayes δ_{BU} est minimax et π_{BU} est la loi a priori la moins favorable.

(2) En particulier, si $m \leq 1/\sqrt{2}$ alors δ_{BU} est minimax et π_{BU} est la moins favorable.

Démonstration. (1) Il faut noter que nous sommes dans les conditions du théorème précédent. Alors, pour $mE_0[h(m,R)] \leq 1/2$ et $m \leq 1$, la fonction de risque de l'estimateur δ_{BU} est croissante en λ. Ce qui veut dire que si $mE_0[h(m,R)] \leq 1/2$ et $m \leq 1$ alors la fonction de risque de l'estimateur δ_{BU} atteint son maximum en $-m$ et m.

En utilisant alors le théorème 3.3.8 de l'appendice, π_{BU} est la loi a priori la moins favorable si $mE_0[h(m,R)] \leq 1/2$ et $m \leq 1$. Mais, la fonction $t(m) = mE_0[h(m,R)] - 1/2$ est croissante en m avec $\lim_{m\to 0} t(m) < 0$ et $\lim_{m\to\sqrt{n}} t(m) > 0$. Ceci nous donne le résultat.

(2) En utilisant le lemme 2.1.1, nous avons $\rho_n(m,r) \leq mr$, pour tout $r \in \Re_n$, donc $mE_0[h(m,R)] \leq m^2$. Si $m^2 \leq 1/2$, nous pouvons alors appliquer la partie (1).

$\square$

Lemme 2.4.1. *Nous avons le résultat asymptotique suivant :*

$$\lim_{n\to\infty} m_1(n) = M_1,$$

où M_1 est solution en m de l'équation $mE[\tanh(mZ)/Z] = 1/2$ où Z admet la distribution normale centrée réduite.

DÉMONSTRATION. Supposons que m et r soient fixés. Nous avons alors :

$$\lim_{n\to\infty} \frac{\sqrt{n}r}{2} \log\left(\frac{1+m/\sqrt{n}}{1-m/\sqrt{n}}\right) = \lim_{n\to\infty} \frac{r}{2} \log\left(1+\frac{1}{\sqrt{n}}\left(\frac{2m}{1-m/\sqrt{n}}\right)\right)^{\sqrt{n}} = mr.$$

De plus, $R = |\sqrt{n}(2\overline{X}-1)|$ et $\sqrt{n}(2\overline{X}-1)$ convergent en loi vers la loi normale centrée réduite. Nous avons alors :

$$\lim_{n\to\infty} E_0\left[\frac{1}{R}\tanh\left(\log\left(\frac{\sqrt{n}R}{2}\frac{1+m/\sqrt{n}}{1-m/\sqrt{n}}\right)\right)\right] = E[\tanh(mZ)/Z].$$

$\square$

Le tableau suivant donne $m_1(n)$ pour certaines valeurs de n.

TAB. 2.2. Valeurs de $m_1(n)$ en fonction de n.

n	1	2	3	4	5	6	7	8	9
$m_1(n)$	0.707	0.816	0.736	0.773	0.744	0.764	0.747	0.760	0.749
n	10	11	12	13	20	50	100	200	∞
$m_1(n)$	0.758	0.750	0.7507	0.751	0.7550	0.7549	0.75495	0.7551	0.7552

Nous proposons par le théorème suivant une condition nécessaire pour que la loi a priori π_{BU} soit la moins favorable. C'est-à-dire que nous aurons une condition de non minimaxité de δ_{BU}.

Théorème 2.4.3. *La condition $m \leq m_2(n)$ avec $m_2(n)$ solution de l'équation $E_0[\rho_n^2(m,R)] + E_m[\rho_n^2(m,R)] - 1 = 0$, est une condition nécessaire pour que la loi a priori π_{BU} soit la moins favorable.*

DÉMONSTRATION. Nous savons que la condition (2.3.1) est nécessaire pour avoir la loi a priori π_{BU} la moins favorable. Cette condition est équivalente à $R(\theta, \delta_{BU}) \leq R(m, \delta_{BU})$ pour tout $-m \leq \theta \leq m$. En particulier la condition $R(0, \delta_{BU}) \leq R(m, \delta_{BU})$ est une condition nécessaire pour avoir π_{BU} la moins favorable. En utilisant la relation (2.4.2) de la fonction de risque de δ_{BU}, cette condition nécessaire est équivalente à :

$$E_0[\rho_n^2(m,R)] + E_m[\rho_n^2(m,R)] - 1 \leq 0.$$

Mais, la fonction $k_n(m) = E_0[\rho_n^2(m,R)] + E_m[\rho_n^2(m,R)] - 1$ est croissante en m avec $\lim_{m \to 0} k_n(m) = -1$. Ceci nous donne le résultat. □

Le tableau suivant nous donne $m_2(n)$ pour quelques valeurs de n.

TAB. 2.3. Valeurs de $m_2(n)$ en fonction de n.

n	1	2	3	4	5	6	7	8	9
m_2	0.707	0.910	0.922	0.963	0.975	0.989	0.9981	1.005	1.0107
n	20	30	40	50	60	70	80	90	100
m_2	1.0355	1.0425	1.0460	1.0481	1.0496	1.0505	1.0514	1.0519	1.0524

Nous considérons à présent la fonction de perte suivante :

$$L^*(\theta, \delta) = q(\lambda^2) L(\theta, \delta) \tag{2.4.3}$$

où q une fonction croissante et strictement positive sur $[0, m^2]$.
Sous cette fonction de perte (2.4.3), l'estimateur de Bayes δ^*_{BU} par rapport à la loi a priori uniforme sur $\Theta(m)$ coïncide avec δ_{BU}.
En effet, nous avons :

$$\delta^*_{BU}(x) = \frac{E_\theta[q(\lambda^2)\theta|x]}{E_\theta[q(\lambda^2)|x]}, \qquad \text{pour tout } x \in \{0, 1, ..., n\}.$$

Comme la loi a priori π_{BU} est concentrée sur $\{-m, m\}$ avec $q((-m)^2) = q(m^2)$ alors $\delta^*_{BU}(x) = E_\theta[\theta|x]$, pour tout $x \in \{0, 1, ..., n\}$.

Soit R^* la fonction de risque associée à L^*, nous avons, pour tout estimateur δ :

$$R^*(\theta, \delta) = q(\lambda^2)R(\theta, \delta), \qquad \text{pour tout } \theta \in \Theta(m).$$

Théorème 2.4.4. *Supposons que $q(t) = (c-t)^{-1}$ pour tout $t \in [0, m^2]$ avec $c > m^2$. Soit g une fonction positive et croissante sur $\Re_n$. Soit la fonction h définie par (2.4.1). Si $E_0[1 - 2mh(R) + m^2g^2(R)/c] \geq 0$ alors $R^*(\theta, \delta_g)$ est croissante en λ, pour $\lambda \in [0, m \wedge 1]$.*

DÉMONSTRATION. Nous avons,

$$4n\frac{\partial}{\partial\lambda}R^*(\lambda, \delta_g) = 4nq(\lambda^2)\frac{\partial}{\partial\lambda}R(\lambda, \delta_g) + 8n\lambda q'(\lambda^2)R(\lambda, \delta_g).$$

En utilisant le théorème 2.4.1, nous avons :

$$4n\frac{\partial}{\partial\lambda}R(\lambda, \delta_g) \geq 2\lambda E_0[1 - 2mh(R)].$$

Similairement à la démonstration du théorème 2.4.1, nous avons :

$$\begin{aligned}
4nR(\lambda, \delta_g) &= 4nE_\lambda[m^2g^2(R) - 2m\lambda g(R)\rho_n(\lambda, R) + \lambda^2], \\
&= E_\lambda[m^2g^2(R) + \lambda^2(1 - 2mh(R))] - 4m\lambda \mathrm{Cov}_\lambda[h(R), R\rho_n(\lambda, R)], \\
&\geq 2E_0[m^2g^2(R) + \lambda^2(1 - 2mh(R))].
\end{aligned}$$

Comme q et q' sont strictement positives, alors nous avons le résultat suivant :

$$\begin{aligned}
4n\frac{\partial}{\partial\lambda}R^*(\lambda, \delta_g) &\geq 2\lambda q(\lambda^2)E_0[1 - 2mh(R)] + 2\lambda q'(\lambda^2)E_0[m^2g^2(R) + \lambda^2(1 - 2mh(R))], \\
&= 2\lambda E_0[(q(\lambda^2) + \lambda^2q'(\lambda^2))(1 - 2mh(R)) + m^2q'(\lambda^2)g^2(R)], \\
&= 2\lambda q'(\lambda^2)E_0[c(1 - 2mh(R)) + m^2g^2(R)], \\
&\geq 0.
\end{aligned}$$

La dernière égalité vient du fait que $q(t) = (c-t)^{-1}$ est solution de l'équation différentielle $q(t) + (t - c)q'(t) = 0$. □

Théorème 2.4.5. *Soit $h(m, r) = \rho_n(m, r)/r$, pour tout $r \in \Re_n$, $r \neq 0$ et $h(m, 0) = g(2/\sqrt{n})/(2/\sqrt{n})$ si $r \in \Re_n$ et $r = 0$. Supposons que $q(t) = (c-t)^{-1}$ pour tout $t \in [0, m^2]$ avec $c > m^2$. Pour $E_0[c(1 - 2mh(R)) + m^2g^2(R)] \geq 0$ et*

$m \leq 1$, la loi a priori π_{BU} est la moins favorable et l'estimateur de Bayes δ_{BU} est minimax.

DÉMONSTRATION. La démonstration est similaire à celle du théorème 2.4.2 $\square$

Les évaluations numériques faites pour la recherche de racines pour la fonction $w_n(m) = E_0[c(1 - 2mh(R)) + m^2 g^2(R)]$, montrent que nous avons pour tout $1 \leq n \leq 235$, $w_n(m) \geq 0$ si et seulement si $m \leq m_3(n)$, avec $m_3(n)$ solution de l'équation $w_n(m) = 0$.

Remarque 2.4.1. *Il faut noter qu'en utilisant la perte L^*, que nous avons une classe d'estimateurs minimax qui couvre des espaces paramétrique plus grands.*

Certaines valeurs de $m_3(n)$ sont données par le tableau suivant.

TAB. 2.4. Valeurs de $m_3(n)$ en fonction de n, pour $c = 1$.

n	1	2	3	4	5	6	7	8	9
$m_3(n)$	0.814	0.934	0.824	0.8655	0.830	0.851	0.833	0.845	0.834
n	10	11	12	13	20	50	100	200	235
$m_3(n)$	0.842	0.835	0.841	0.836	0.839	0.833	0.8385	0.8387	0.8389

Chapitre 3

PROBLÈMES D'ESTIMATION MINIMAX AVEC RESTRICTION SUR L'ESPACE DES PARAMÈTRES

3.1. Introduction

Ce chapitre traite les problèmes d'estimation minimax avec une restriction sur l'espace des paramètres. Beaucoup de chercheurs ont contribué à ce problème. Par exemple, DasGupta (1985) étudie le problème général d'un vecteur de paramètres lorsque l'espace paramétrique est convexe, sous la fonction de perte quadratique, pour une famille générale de distributions. Il montre, sous certaines conditions, qu'il existe une loi a priori la moins favorable concentrée sur la frontière de l'espace paramétrique restreint Θ, lorsque Θ est assez petit. Kempthorne (1987) montre, sous certaines conditions (voir appendice), qu'une loi a priori la moins favorable π_0 existe et que soit la fonction de risque de l'estimateur de Bayes par rapport à π_0 est constante ou soit le support de π_0 est fini et discret. Il considère les problèmes auxquels les lois a priori les moins favorables ont leur support fini et discret et propose un algorithme qui permet de construire ces lois a priori les moins favorables.

Dans le cas d'une distribution continue, Casella et Strawderman (1981) traitent le cas normal $N(\theta, \sigma)$, avec $a \leq \theta \leq b$ et σ connu. Ils ont montré que pour $(b-a)/2 \leq m_0\sigma$, avec $m_0 \approx 1.0567$, la loi a priori concentrée sur $\{a, b\}$ est la moins favorable et l'estimateur de Bayes associé est minimax. Bickel (1981) considère le même problème que ces derniers et proposent une solution asymptotique. Bader et Bischoff (2003) considèrent le problème d'estimation d'un paramètre lorsque ce dernier appartient à un intervalle $[\theta_0, \theta_0 + m]$ où $0 < m \leq M$ avec θ_0, m et M connus pour une fonction de perte strictement convexe. Ils montrent que, sous certaines conditions sur la distribution et sur la fonction de perte, l'estimateur de Bayes associé à la loi a priori concentrée sur $\{\theta_0, \theta_0+m\}$ est minimax si m est assez petit. Eichenauer (1986) et Berry (1993) s'intéressent au cas d'une distribution exponentielle pour la fonction de perte quadratique, lorsque le paramètre de position θ appartient à un intervalle $[0, m]$, m connu. Ils montrent que la condition $m \leq m_0 \approx 0.913$ est une condition nécessaire et suffisante pour avoir la loi a priori la moins favorable concentrée sur les deux points $\{0, m\}$ et l'estimateur de Bayes par rapport à cette loi a priori minimax. Pour le problème d'une famille de distributions tronquées ayant la densité de la forme $f(x|\theta) = c(\theta) f_0(x) I_{[\theta,\theta+1]}(x)$, pour $x \in I\!R$, $\theta \in [0, m]$ et $c(\theta)$ la constante de normalisation, Eichenauer et Fieger (1992) montrent l'existence d'un m_0 tel que pour tout $m \leq m_0$, la loi a priori la moins favorable soit concentrée sur les valeurs $\{0, m\}$ et l'estimateur de Bayes correspondant est minimax pour une fonction de perte convexe. Cependant, ils ne donnent la valeur de m_0 que pour le cas particulier d'une perte absolue et d'une distribution uniforme ($f_0 \equiv 1$).

Dans le cas d'une distribution discrète, Moors (1985) traite l'estimation minimax du paramètre θ, pour une loi Binomiale(n, θ), lorsque θ appartient à un intervalle symétrique. Il montre que l'estimateur de Bayes dépend de la loi a priori correspondante seulement par un nombre fini de ses moments. Il montre aussi que le nombre maximal de points dans le support de la loi a priori la moins favorable ne doit pas excéder $2t+1$ où t est le plus grand entier inférieur où égale à $(n+3)/4$. Il donne également la liste de tous les estimateurs de Bayes minimax accompagnés des lois a priori les moins favorables pour $1 \leq n \leq 3$ selon la valeur de t. Berry (1989) utilise une approche plus directe que Moors (1985) en montrant, pour $n = 1, 2, 3, 4$, que la loi a priori la moins favorable est supportée par deux ou trois points selon les valeurs sur les bornes de p. Marchand et MacGibbon (2000) étudient le problème d'estimation minimax et linéaire minimax du paramètre θ d'une distribution Binomiale(n, θ), sous la fonction de perte quadratique et fonction de perte normalisée, lorsque le paramètre θ appartient à l'intervalle $[a, b]$ où a et b sont connus. Ils proposent des résultats analytiques pour le cas où $a = 0$ et b est petit, et aussi des résultats numériques pour le cas général. Johnstone et MacGibbon (1992) traitent le problème d'estimation minimax de la moyenne θ d'une distribution de Poisson lorsque $\theta \in [0, m]$, m connu, sous la fonction de perte normalisée. Ils montrent que, lorsque $0 < m \leq m_0(1) \approx 0.57$, la loi a priori la moins favorable est concentrée sur le point $\{m\}$ et lorsque $m_0(1) < m \leq m_0(2) \approx 1.27$, la loi priori la moins favorable est concentrée sur les deux points $\{0, m\}$.

En général, ces auteurs cherchent à démontrer la minimaxité de l'estimateur de Bayes par rapport à une loi a priori concentrée sur la frontière de l'espace paramétrique. Mais, la plupart de leurs résultats se réfèrent seulement à des pertes symétriques et le plus souvent à la perte quadratique. Les pertes symétriques, telles la perte quadratique et la perte L^q avec $q \geq 1$, imposent une même pénalité sur la surestimation et la sous-estimation dans la même échelle de grandeur. Mais, il y a de nombreux exemples en pratique où il est désirable de pénaliser la surestimation et la sous-estimation de façons différentes. Une option utile parmi les pertes asymétriques est la fonction de perte LINEX. Elle a été introduite par Varian (1975) et plusieurs développements ont été obtenus par Zellner (1986). Il faut noter que cette fonction de perte est non seulement asymétrique mais aussi convexe. Elle porte le nom LINEX du fait qu'elle croît presque linéairement d'un côté de zéro et croît presque exponentiellement de l'autre côté. Lorsque γ est très proche de zéro, cette fonction de perte est presque symétrique et se rapproche de la fonction de perte quadratique. Lorsque γ est positif (respectivement négatif), la surestimation (respectivement la sous-estimation) est considérée plus grave que la sous-estimation (respectivement la surestimation). Récemment, plusieurs travaux dont Bischoff, Fieger et Wulfert (1995), Khattree (1992), Wan, Zou et Lee

(2000) et Parsian et Kirmani (2002) ont porté sur l'estimation avec la perte LINEX. Parsian et Kirmani (2002) ont fourni une revue détaillée des travaux reliés à cette fonction de perte.

Nous allons étudier alors, dans ce chapitre, le problème d'estimation minimax d'un paramètre $\theta \in \theta(m) = [0, m]$, $m > 0$, sous les fonctions de perte L^2 et LINEX. Par exemple Wan, Zou et Lee (2000) considèrent le cas d'une distribution de Poisson pour une perte LINEX. Ils utilisent la technique de convexité pour donner des conditions suffisantes sur m pour que la loi a priori la moins favorable soit concentrée sur $\{0, m\}$ et que l'estimateur de Bayes correspondant soit minimax. Bischoff, Fieger et Wulfert (1995) ont montré dans le cas d'une distribution normale $N(\theta, \sigma^2)$ et d'une perte LINEX, lorsque σ est connu et $\theta \in [-m, m]$, que la loi a priori la moins favorable est concentrée sur $\{-m, m\}$ et l'estimateur de Bayes correspondant est minimax, pour m assez petit . Habituellement, ces auteurs proposent des conditions sur m pour que la loi a priori la moins favorable soit concentrée sur la frontière de l'espace paramétrique restreint et l'estimateur de Bayes correspondant soit minimax. Mais le plus souvent, ces conditions ne sont que suffisantes pour une perte LINEX. Un de nos objectifs est de trouver, pour notre problème, des conditions nécessaires et suffisantes sur m pour que la loi a priori la moins favorable soit concentrée sur $\{0, m\}$ et que l'estimateur de Bayes correspondant soit minimax. Certaines de nos conditions sont en effet nécessaires et suffisantes, tandis que d'autres ne sont que suffisantes. Nos conditions sont de la forme $0 < m \leq m_0$. Nous proposerons des équations explicites et des évaluations numériques pour trouver m_0, ainsi que des bornes inférieures et supérieures de m_0.

D'abord, dans le cas d'une distribution continue, nous considérons la distribution donnée par la densité $f(x|\theta) = f_0(x - \theta)I_{[\theta,\infty)}(x)$, pour $0 \leq \theta \leq m$. Pour ce faire, nous allons proposer des propriétés générales pour une fonction de perte strictement convexe. Ainsi, nous utilisons ces propriétés générales pour proposer des solutions minimax dans les cas particuliers d'une distribution exponentielle sous la fonction de perte LINEX et d'une distribution uniforme pour les fonctions de perte quadratique et LINEX. Mais, dans le cas d'une distribution exponentielle avec la perte LINEX lorsque le paramètre de la perte γ est positif et dans le cas d'une distribution uniforme avec les pertes quadratique et LINEX, nous proposerons des conditions nécessaires et suffisantes pour lesquelles la loi a priori la moins favorable soit concentrée sur $\{0, m\}$. Nous donnerons également des conditions suffisantes dans le cas d'une distribution exponentielle avec la perte LINEX lorsque γ est négatif.

Ensuite, pour une distribution discrète, nous allons considérer une fonction de masse donnée par p_θ de support $\{0, 1, 2, ..., s\}$, $s < \infty$ ou $\{0, 1, 2, ...\}$. Nous proposerons des propriétés et résultats généraux sous une fonction de perte strictement convexe ρ. Plus particulièrement, nous allons donner explicitement des conditions (en général nécessaires et suffisantes) sur m pour lesquelles la loi a priori la moins favorable π_m soit concentrée sur la frontière de l'espace paramétrique restreint $\Theta(m)$. D'abord, dans le cas d'une perte quadratique, nous allons poser des conditions sur la fonction de masse (ces conditions sont satisfaites par plusieurs distributions dont la distribution de Poisson, la distribution Binomiale et la distribution Binomiale négative) et nous donnerons les valeurs de m_0. Ces résultats viennent des travaux en cours de Marchand et Parsian (2002). Nous appliquerons ces résultats dans le cas d'une distribution de Poisson, d'une distribution Binomiale et d'une distribution Binomiale négative. Ensuite pour une perte LINEX et une distribution de Poisson, nous dégagerons des conditions nécessaires et suffisantes sur m, dans le cas où le paramètre γ de la perte est négatif, pour lesquelles π_m soit la loi a priori la moins favorable. Nous proposerons également, pour $\gamma < 0$, des conditions nécessaires et suffisantes pour lesquelles la fonction de risque de l'estimateur de Bayes δ_m associé à π_m soit convexe, ce qui mène à des conditions suffisantes pour la minimaxité de δ_m.

La recherche de solutions minimax pour nos problèmes est surtout basée sur des conditions de Kempthorne (1987), des théorèmes et lemmes proposés dans l'appendice.

3.2. Distribution continue

Dans cette section, nous nous intéressons au problème d'estimation minimax du paramètre de position $\theta \in \Theta(m)$, $m > 0$, d'une famille de distribution de la forme $f(x|\theta) = f_0(x-\theta)I_{[\theta,+\infty)}(x)$, sous une fonction de perte strictement convexe. En développant des propriétés générales pour les estimateurs bayésiens associés à des lois a priori concentrées sur la frontière $\{0, m\}$, nous trouvons des estimateurs minimax pour une distribution exponentielle et une distribution uniforme lorsque $\theta \in \Theta(m)$. Plus précisément, nous recherchons des conditions sur m de la forme $m \leq m_0$ telle que la loi a priori la moins favorable soit concentrée sur la frontière de l'espace paramétrique restreint dans les cas particuliers d'une distribution exponentielle sous la fonction de perte LINEX et d'une distribution uniforme sous les fonctions de perte quadratique et LINEX.

3.2.1. Préliminaires, notations et propriétés générales

Soit X une variable aléatoire issue d'une famille de densités qui ont la forme :

$$f(x|\theta) = f_0(x-\theta)I_{[\theta,+\infty)}(x). \tag{3.2.1}$$

Considérons des fonctions de perte invariantes pour les translations et qui sont la forme suivante :

$$L(\theta, d) = \rho(d-\theta), \tag{3.2.2}$$

avec ρ strictement convexe, $\rho(x) \geq 0$ pour tout $x \in I\!R$ et $\rho(x) = 0$ si et seulement si $x = 0$.

Exemple 3.2.1. *Exemples de fonctions strictement convexes*

(1) Pour une perte L^q, nous avons $\rho(x) = |x|^q, \quad q > 1$.

(2) Pour une perte LINEX, nous avons :

$$\rho(x) = \beta\{\exp(\gamma x) - \gamma x - 1\}, \tag{3.2.3}$$

où $\gamma \neq 0$ est le paramètre de forme et $\beta > 0$ est un facteur de proportionnalité. Sans perte de généralité, nous choisissons $\beta = 1$.

On considère la loi a priori concentrée sur les points $\{0, m\}$, donnée par :

$$\pi_\alpha(0) = 1 - \pi_\alpha(m) = \alpha. \tag{3.2.4}$$

La loi a posteriori est également concentrée sur la frontière $\{0, m\}$ et elle est donnée par $\pi_\alpha(0|x) = 1 - \pi_\alpha(m|x)$ avec

$$\pi_\alpha(0|x) = \begin{cases} 1 & \text{si } 0 \le x < m, \\ k_\alpha(x, m) & \text{si } m \le x < \infty, \end{cases}$$

où

$$k_\alpha(x, m) = \alpha f_0(x)/(\alpha f_0(x) + (1 - \alpha) f_0(x - m)). \tag{3.2.5}$$

Il s'en suit que la perte a posteriori est égale à :

$$E[\rho(d - \theta)|x] = \begin{cases} \rho(d) \text{ si } 0 \le x < m, \\ \rho(d) k_\alpha(x, m) + \rho(d - m)(1 - k_\alpha(x, m)), \text{ si } m \le x < \infty. \end{cases} \tag{3.2.6}$$

Par conséquent, la loi a priori concentrée sur les deux points $\{0, m\}$ donne un estimateur de Bayes de la forme :

$$\delta_{\alpha,m}(x) = \begin{cases} 0 & \text{si } 0 \le x < m, \\ y_{\alpha,m}(x) & \text{si } m \le x < \infty, \end{cases} \tag{3.2.7}$$

où $y_{\alpha,m}(x)$ minimise en d l'expression suivante :

$$\rho(d) k_\alpha(x, m) + \rho(d - m)(1 - k_\alpha(x, m)). \tag{3.2.8}$$

Nous proposons, par l'exemple suivant, des expressions générales pour ces estimateurs de Bayes $\delta_{\alpha,m}$ pour les fonctions de perte L^q, $q > 1$ et LINEX.

Exemple 3.2.2. *(1) Pour la perte L^q avec $q > 1$, nous avons*

$$y_{\alpha,m}(x) = m \frac{[(1 - \alpha) f_0(x - m)]^{1/(q-1)}}{[(1 - \alpha) f_0(x - m)]^{1/(q-1)} + [\alpha f_0(x)]^{1/(q-1)}}.$$

(2) Pour la perte LINEX (3.2.3), nous avons :

$$y_{\alpha,m}(x) = \frac{1}{\gamma} \log \left(\frac{\alpha f_0(x) + (1 - \alpha) f_0(x - m)}{\alpha f_0(x) + \exp(-\gamma m)(1 - \alpha) f_0(x - m)} \right). \tag{3.2.9}$$

Pour obtenir les résultats de l'exemple précédent, nous avons calculé, pour $\rho(x) = |x|^q$ et $\rho(x) = \exp(\gamma x) - \gamma x - 1$, la dérivée première de l'expression (3.2.8) en d et nous avons cherché la valeur de d pour laquelle cette dérivée est nulle. La convexité de ρ nous assure que nous obtenons bien un minimum.

Nous mesurons la performance de l'estimateur de Bayes $\delta_{\alpha,m}$ par la fonction de risque de cet estimateur. Cette fonction de risque est donnée par :

$$\begin{aligned} R(\theta, \delta_{\alpha,m}) &= E_\theta[\rho(\delta_{\alpha,m}(X) - \theta)], \\ &= \rho(-\theta) \int_\theta^m f(x-\theta)dx + \int_m^\infty \rho(y_{\alpha,m}(x) - \theta) f(x-\theta)dx, \ \forall\, \theta \in \Theta(m). \end{aligned} \tag{3.2.10}$$

Nous proposons à présent un lemme qui nous sera utile par la suite. if faut noter que ce lemme est général et qu'il ne se limite pas seulement aux distributions de la forme (3.2.1).

Lemme 3.2.1. *Le risque minimax* $\overline{R}(m) = \inf_\delta \sup_{0 \le \theta \le m} R(\theta, \delta)$ *est croissant en* m.

Démonstration. Soient m_1 et m_2 avec $0 < m_1 < m_2$. Nous avons :

$$\sup_{0 \le \theta \le m_1} R(\theta, \delta) \le \sup_{0 \le \theta \le m_2} R(\theta, \delta) \text{ pour tout } \delta.$$

Ceci nous donne le résultat :

$$\inf_\delta \sup_{0 \le \theta \le m_1} R(\theta, \delta) \le \inf_\delta \sup_{0 \le \theta \le m_2} R(\theta, \delta).$$

□

3.2.2. Distribution exponentielle

Dans cette partie, nous allons étudier le cas d'une distribution exponentielle de densité donnée par :

$$f(x|\theta) = \frac{1}{\sigma} \exp\left(-\frac{x-\theta}{\sigma} \right) I_{[\theta,\infty]}(x),$$

$\theta \in [0, m]$ et $\sigma > 0$ connu.

Dans le cadre de la condition (3.2.1), nous avons $f_0(x) = \sigma^{-1} \exp(-x/\sigma) I_{[\theta,\infty]}(x)$. Mais, on peut sans perte de généralité supposer qu'on a affaire à une seule observation X de fonction de densité $f(x|\theta) = \exp(-(x-\theta)) I_{[\theta,\infty]}(x)$. En effet, la statistique exhaustive $T(X_1, X_2, ..., X_n) = (n/\sigma)\min(X_1, X_2, ...X_n)$ admet la densité :

$$f(t|\theta') = \exp(-(t-\theta'))I_{[\theta',\infty]}(t),$$

où $\theta' = n\theta/\sigma$ et la contrainte $0 \leq \theta \leq m$ se traduit par $0 \leq \theta' \leq m' = nm/\sigma$.

Ainsi, pour la loi exponentielle et par la relation (3.2.5), k_α prend la forme suivante :

$$k_\alpha(m, x) = \frac{\alpha}{\alpha + \exp(m)(1-\alpha)}, \quad \text{pour tout } x \geq \theta.$$

Il faut remarquer que sur les régions $[0, m)$ et $[m, \infty)$, cette perte a posteriori ne dépend pas de x. Ce qui simplifie la recherche des estimateurs de Bayes .

Nous sommes maintenant à la recherche de conditions pour lesquelles la loi a priori concentrée sur $\{0, m\}$ soit la moins favorable. Mais, il faut noter que nous vérifions les conditions A à E de l'appendice. En utilisant alors le théorème 3.3.5 (appendice) et le lemme 3.3.4 (appendice), il est nécessaire d'avoir :

$$R(0, \delta_{\alpha,m}) = R(m, \delta_{\alpha,m}), \tag{3.2.11}$$

pour qu'une loi a priori concentrée sur $\{0, m\}$ soit la moins favorable. Nous appelons "égalisateur" un estimateur $\delta_{\alpha,m}$ qui vérifie l'équation (3.2.11).
Le prochain lemme identifie un unique "égalisateur" pour chaque valeur de $m > 0$ fixée.

Lemme 3.2.2. *Parmi les estimateurs de la forme (3.2.7), il existe exactement un "égalisateur" donné par :*

$$\delta_m(x) = \begin{cases} 0 \text{ si } 0 \leq x < m \\ y(m) \text{ si } m \leq x < \infty, \end{cases}$$

où $y(m)$ satisfait l'équation suivante en y :

$$\exp(-m)\rho(y) = \rho(y-m). \tag{3.2.12}$$

Nous notons par π_m la loi a priori par rapport à l'estimateur de Bayes δ_m.

DÉMONSTRATION. Par la relation (3.2.10), la fonction de risque de $\delta_{\alpha,m}$ devient :

$$R(\theta,\delta_{\alpha,m}) = \rho(-\theta)[1-\exp(\theta-m)] + \exp(\theta-m)\rho(y_{\alpha,m}-\theta). \tag{3.2.13}$$

Nous avons alors,

$$R(0,\delta_{\alpha,m}) = \exp(-m)\rho(y_{\alpha,m}(m)) \text{ et } R(m,\delta_{\alpha,m}) = \rho(y_{\alpha,m}(m)-m).$$

Ainsi, la relation (3.2.11) est équivalente à la relation suivante :

$$\exp(-m)\rho(y_{\alpha,m}) = \rho(y_{\alpha,m}-m).$$

Puisque $y_{\alpha,m}$ est une fonction continue en α lorsque α varie dans $[0,1]$ avec $y_{0,m} = m$ et $y_{m,m} = 0$ alors la relation précédente est équivalente à la relation (3.2.12). Considérons la fonction $\Delta(y) = \exp(-m)\rho(y) - \rho(y-m)$, pour $y \in [0,m]$. Cette fonction est continue et strictement croissante en y (puisque $\rho(z)$ est croissant pour $z > 0$ et décroissant sur $z < 0$) avec $\Delta(0) = -\rho(-m) \leq 0$ et $\Delta(m) = \exp(-m)\rho(m) \geq 0$. Il existe alors un unique $y \in [0,m]$ tel que $\Delta(y) = 0$. Ce qui achève la démonstration. □

Exemple 3.2.3. *Pour la perte L^q, $q \geq 1$, la relation (3.2.12) devient :*

$$\exp(-m)y^q = (m-y)^q.$$

Nous avons alors à résoudre une équation en y. La solution de cette équation est donnée par :

$$y(m) = m\frac{\exp(m/q)}{\exp(m/q)+1}, \quad m \geq 0.$$

Quoiqu'il en n'est pas nécessaire, nous pouvons déduire de cette relation et de la partie (1) de l'exemple 3.2.2 la valeur de α correspondante à $y(m)$ comme étant égale à :

$$\alpha_m = \frac{\exp(m/q)}{1+\exp(m/q)} = \frac{y(m)}{m}, \quad m > 0.$$

Nous savons par Bader et Bischoff (2003) que, pour une large classe de distributions et pour une fonction de perte strictement convexe, il existe m_0 tel que pour tout $m \leq m_0$, la loi a priori concentrée sur $\{0, m\}$ est la moins favorable. Nous allons utiliser ce résultat pour donner quelques propriétés de $y(m)$ par le prochain lemme. Ces propriétés nous seront utiles par la suite.

Lemme 3.2.3. *Si, pour $m \leq m_0$, la loi a priori concentrée sur $\{0, m\}$ est la moins favorable alors*

(1) la fonction $y(m)$ est croissante en m,

(2) $y'(m) \leq 1$.

DÉMONSTRATION. Nous savons que si la loi a priori π_m est la moins favorable, alors nous aurons les égalités suivantes pour le risque minimax $\overline{R}(m)$:

$$\overline{R}(m) = R(0, \delta_m) = R(m, \delta_m). \tag{3.2.14}$$

Cette relation est équivalente à :

$$\overline{R}(m) = \exp(-m)\rho(y(m)) = \rho(y(m) - m). \tag{3.2.15}$$

(1) D'une part, pour $m \leq m_0$, en dérivant le risque minimax par rapport à m et en utilisant le lemme (3.2.1), nous aurons :

$$\frac{\partial}{\partial m}\overline{R}(m) = \exp(-m)[-\rho(y(m)) + y'(m)\rho'(y(m))] \geq 0.$$

Les propriétés de la fonction de perte ρ nous assure que $\rho'(y(m)) > 0$ et nous devons avoir nécessairement $y'(m) \geq 0$.

(2) D'autre part, pour $m \leq m_0$, en dérivant le risque minimax, donnée à droite de (3.2.15), par rapport à m et en utilisant le lemme (3.2.1), nous aurons :

$$[y'(m) - 1]\rho'(y(m) - m) \geq 0$$

Ce qui nous donne bien $y'(m) - 1 \leq 0$ car $y(m) \leq m$ et $\rho'(z) < 0$ pour $z < 0$. Ce qui achève la démonstration. $\square$

Nous voulons à présent trouver des conditions nécessaires pour lesquelles la loi a priori π_m soit la moins favorable et l'estimateur de Bayes δ_m soit minimax.

Mais, nous savons qu'un critère pour que la loi a priori π_m soit la moins favorable est :

$$\sup_{0\leq\theta\leq m} R(\theta,\delta_m) = \max\{R(0,\delta_m), R(m,\delta_m)\} = R(0,\delta_m) = R(m,\delta_m). \quad (3.2.16)$$

Cependant, si la condition (3.2.16) est vraie alors la fonction de risque de δ_m est décroissante en $\theta = 0$ et elle est croissante en $\theta = m$. Ces deux conditions se traduisent par les résultats du théorème suivant.

Théorème 3.2.1. *Des conditions nécessaires pour que la relation (3.2.16) soit vraie (donc π_m soit la moins favorable) sont :*

(1)

$$\rho(y(m)) \leq \rho'(y(m)), \quad (3.2.17)$$

(2)

$$\rho'(y(m)-m) \leq \rho(y(m)-m) - \rho(-m). \quad (3.2.18)$$

Démonstration. Par la relation (3.2.13) du risque de δ_m, la dérivée de cette fonction de risque est donnée par :

$$\frac{\partial}{\partial\theta}R(\theta,\delta_m) = -\rho'(-\theta) + \exp(\theta-m)[\rho(y(m)-\theta) - \rho(-\theta) - \rho'(y(m)-\theta) + \rho'(-\theta)].$$

Nous en déduisons alors $\frac{\partial}{\partial\theta}R(\theta,\delta_m)\big|_{\theta=0} = \exp(-m)[\rho(y(m)) - \rho'(y(m))]$ et $\frac{\partial}{\partial\theta}R(\theta,\delta_m)\big|_{\theta=m} = \rho(y(m)-m)-\rho(-m)-\rho'(y(m)-m)$. Les conditions nécessaires $\frac{\partial}{\partial\theta}R(\theta,\delta_m)\big|_{\theta=0} \leq 0$ et $\frac{\partial}{\partial\theta}R(\theta,\delta_m)\big|_{\theta=m} \geq 0$ sont alors équivalentes aux conditions (3.2.17) et (3.2.18) respectivement. $\square$

Dans le cas d'une perte LINEX, la condition (3.2.17) devient une condition très simple de la forme $m \leq m_*(\gamma)$. Le prochain corollaire nous procure ce résultat.

Corollaire 3.2.1. *Pour une perte LINEX (3.2.3), nous avons les résultats suivants :*

(1) *si $\gamma \geq 1$, la condition nécessaire (3.2.17) se réalise pour tout $m > 0$,*

(2) *si $\gamma < 1$, cette condition se réalise si et seulement si $m \leq m_*(\gamma)$ avec $m_*(\gamma)$ vérifiant l'équation $h_\gamma(m) = 0$ où $h_\gamma(m) = (1-\gamma)(\exp(\gamma y(m)) - 1) - \gamma y(m)$.*

DÉMONSTRATION. Sous la fonction de perte LINEX, la condition (3.2.17) devient :

$$(1-\gamma)[\exp \gamma y(m) - 1] - \gamma y(m) \leq 0, \tag{3.2.19}$$

c'est-à-dire $h_\gamma(m) \leq 0$. Il faut d'abord remarquer que $y(m)$ est une fonction continue en m avec $y(0) = 0$. Ce qui implique que $h_\gamma(m)$ est également une fonction continue en m avec $h_\gamma(0) = 0$.

• Pour $\gamma \geq 1$, il est facile de voir que la fonction $h_\gamma(m)$ est négative pour tout $m \geq 0$.

• Pour $\gamma < 1$, la dérivée de la fonction $h_\gamma(m)$ par rapport à m est donnée par $h'_\gamma(m) = \gamma y'(m)[(1-\gamma)\exp(\gamma y(m)) - 1]$, pour tout $m \geq 0$. Par le lemme 3.2.3, nous avons $y'(m) \geq 0$, donc la fonction $h'(m)$ a le même signe que la fonction $k_\gamma(m) = \gamma[(1-\gamma)\exp(\gamma y(m)) - 1]$. Mais, la fonction $k_\gamma(m)$ est croissante en m du fait que $y(m)$ est croissant en m, de plus $k_\gamma(0) = -\gamma^2 < 0$. Donc la fonction k_γ fait au plus un changement de signe en passant du négatif au positif pour $m > 0$. La fonction h_γ est alors soit décroissante ou soit décroissante et ensuite croissante. Mais, il est facile de voir avec la relation (3.2.9) que $\lim_{m\to\infty} y(m) = +\infty$. Il s'en suit alors que $\lim_{m\to\infty} h_\gamma(m) > 0$. En utilisant le fait que $h_\gamma(0) = 0$, nous avons $h_\gamma(m)$ qui passe du négatif au positif. Ceci nous donne le résultat. □

Nous allons travailler, dans tout ce qui suit de cette sous-section, avec la fonction de perte LINEX. Sous cette fonction de perte, nous allons proposer aux théorèmes 3.2.2 et 3.2.3 respectivement :

(A) des conditions nécessaires et suffisantes pour que la loi a priori π_m soit la moins favorable dans le cas de la perte LINEX avec paramètre de la perte γ positif. Mais, les évaluations numériques montrent que ces conditions nécessaires et suffisantes sont de la forme $m \leq m_0(\gamma)$.

(B) des conditions suffisantes de la forme $m \leq m_1(\gamma)$ pour que la loi a priori π_m soit la moins favorable pour une perte LINEX avec γ négatif.

Nous obtenons nos conditions nécessaires et suffisantes en montrant que la dérivée troisième par rapport à θ de la fonction de risque de δ_m est négative. Ce comportement de la troisième dérivée, conjugué à la contrainte

$R(0,\delta_m) = R(m,\delta_m)$ nous permet de travailler avec la condition $\frac{\partial}{\partial\theta}R(\theta,\delta_m)\big|_{\theta=m} \geq 0$ pour que la loi a priori π_m soit la moins favorable. Alors cette dernière condition devient nécessaire et suffisante. Nos conditions suffisantes sont obtenues en s'appuyant sur la technique de la convexité. C'est-à-dire, que nous montrons que ces conditions sont nécessaires et suffisantes pour que la fonction de risque de δ_m soit convexe.

Nous commençons alors par donner la fonction de risque de δ_m. Par la relation (3.2.13) du risque de δ_m, elle devient :

$$R(\theta,\delta_m) = (\exp(-\gamma\theta) + \gamma\theta - 1) + \exp(\theta - m)[\exp(\gamma(y(m) - \theta) - \exp(-\gamma\theta) - \gamma y(m)]. \tag{3.2.20}$$

En utilisant alors cette fonction de risque de δ_m et ses dérivées premières, nous proposons le résultat (B) par le prochain théorème.

Théorème 3.2.2. *Pour tout $\gamma \neq 0$, nous avons les résultats suivants.*

(1) La fonction de risque de δ_m est convexe en θ si et seulement si $A_\gamma(m) \geq 0$ avec $A_\gamma(m) = \gamma^2 + (1-\gamma)^2(\exp(\gamma y(m)) - 1) - \gamma y(m)\exp(\gamma m)$. De plus, lorsque $\gamma < 0$, la condition $A_\gamma(m) \geq 0$ est équivalente à $m \leq m_1(\gamma)$ où $m_1(\gamma)$ est solution de l'équation $A_\gamma(m) = 0$ en m.

(2) Lorsque $\gamma < 0$, l'estimateur de Bayes δ_m est minimax et la loi a priori π_m est la moins favorable si $m \leq m_1(\gamma)$.

DÉMONSTRATION. (1) Nous avons, par la relation (3.2.20), la dérivée deuxième de la fonction de risque de δ_m qui est donnée par :

$$\begin{aligned}\frac{\partial^2}{\partial\theta^2}R(\theta,\delta_m) &= \exp(-\gamma\theta)\{\gamma^2 + \exp(\theta - m)[(1-\gamma)^2(\exp(\gamma y(m)) - 1) - \gamma y(m)\exp(\gamma\theta)]\}, \\ &= \exp(-\gamma\theta)A_\gamma(\theta).\end{aligned} \tag{3.2.21}$$

avec $A_\gamma(\theta) = \gamma^2 + \exp(\theta - m)B_\gamma(\theta)$ et
$B_\gamma(\theta) = (1-\gamma)^2[\exp(\gamma y(m)) - 1] - \gamma y(m)\exp(\gamma\theta)$.
Mais, pour $\gamma \neq 0$, $B_\gamma(\theta)$ est décroissante en θ. Nous avons alors $B_\gamma(\theta) \geq B_\gamma(m)$, pour tout $\gamma \neq 0$. Ceci nous donne $A_\gamma(\theta) \geq \gamma^2 + \exp(\theta - m)B_\gamma(m)$, pour tout $\gamma \neq 0$. Nous avons deux possibilités.

- Soit $B_\gamma(m)$ est positif alors dans ce cas, nous avons $A_\gamma(\theta)$ est positif, pour tout $\theta \in \Theta(m)$.
- Soit $B_\gamma(m)$ est négatif alors dans ce cas, $A_\gamma(\theta) \geq A_\gamma(m)$, pour tout $\theta \in \Theta(m)$.

Nous avons alors $A_\gamma(\theta) \geq 0$ pour tout $\theta \in [0, m]$ si et seulement si $A_\gamma(m) \geq 0$.
En outre, lorsque $\gamma < 0$, nous avons :

$$\begin{aligned} \frac{\partial}{\partial m} A_\gamma(m) &= \gamma(1-\gamma)^2 y'(m)\exp(\gamma y(m)) - \gamma y'(m)\exp(\gamma m) - \gamma^2 y(m)\exp(\gamma m), \\ &= \gamma y'(m)[(1-\gamma)^2\exp(\gamma y(m)) - \exp(\gamma m)] - \gamma^2 y(m)\exp(\gamma m) \leq 0. \end{aligned}$$

Donc, la fonction $A_\gamma(m)$ est décroissante en m avec $A_\gamma(0) = \gamma^2 > 0$ et $\lim_{m\to\infty} A_\gamma(m) < 0$. Ceci nous donne le résultat.
(2) C'est une conséquence directe du résultat (1) et du théorème 3.3.8 (appendice). □

Pour obtenir le résultat (A) de la discussion, nous avons besoin du lemme suivant.

Lemme 3.2.4. *Sous la fonction de perte LINEX et pour $\gamma > 0$, $\frac{\partial^2}{\partial\theta^2}R(\theta, \delta_m)$ est décroissante sur $\Theta(m)$ pour tout $m \leq m_*(\gamma)$ où $m_*(\gamma)$ est donné par le corollaire 3.2.1.*

Démonstration. Par la relation (3.2.21), la dérivée troisième de la fonction de risque de l'estimateur δ_m est donnée par :

$$\begin{aligned} \frac{\partial^3}{\partial\theta^3}R(\theta, \delta_m) &= \exp(-\gamma\theta)[-\gamma^3 + \exp(\theta - m)[(1-\gamma)^3(\exp(\gamma y(m)) - 1) - \gamma y(m)\exp(\gamma\theta)]], \\ &= \exp(-\gamma\theta)[-\gamma^3 + \exp(\theta - m)H_\gamma(\theta, m)], \end{aligned}$$

où $H_\gamma(\theta, m) = (1-\gamma)^3[\exp(\gamma y(m)) - 1] - \gamma y(m)\exp(\gamma\theta)$.
(1) Si $\gamma \geq 1$, le résultat est immédiat.
(2) Nous supposons que $0 < \gamma < 1$. Mais, comme la fonction $H_\gamma(\theta, m)$ est décroissante en θ alors nous avons :

$$\begin{aligned} H_\gamma(\theta, m) &\leq H_\gamma(0, m), \\ &= (1-\gamma)^3[\exp(\gamma y(m)) - 1] - \gamma y(m), \\ &\leq (1-\gamma)[\exp(\gamma y(m)) - 1] - \gamma y(m). \end{aligned}$$

En utilisant alors la condition nécessaire (3.2.19), nous aurons $H_\gamma(\theta, m) \leq 0$, pour tout $\theta \in \Theta(m)$ et $m \leq m_*(\gamma)$. Ceci nous donne le résultat. $\square$

Théorème 3.2.3. *Pour $\gamma > 0$, la loi a priori π_m est la moins favorable et l'estimateur de Bayes δ_m est minimax si et seulement si $m \in \{m : G_\gamma(m) \geq 0\} \neq \emptyset$, avec $G_\gamma(m) = \gamma + (1-\gamma)\exp(\gamma(y(m) - m)) - \exp(-\gamma m) - \gamma y(m)$.*

DÉMONSTRATION. Par le lemme précédent, nous avons la dérivée troisième de la fonction de risque négative pour tout $m \leq m_*(\gamma)$ et $\gamma \geq 0$. La condition (3.2.18) devient alors une condition nécessaire et suffisante pour que la fonction de risque atteigne son maximum en 0 et m. Mais, cette condition est équivalente à $G_\gamma(m) \geq 0$. Finalement, puisque $\frac{\partial}{\partial m} G_\gamma(m)\big|_{m=0} > 0$ et $G_\gamma(0) = 0$, l'ensemble $\{m : G_\gamma(m) \geq 0\}$ n'est pas vide. En utilisant alors le théorème 3.3.8 (appendice), nous avons le résultat. $\square$

Remarque 3.2.1. *Il est intéressant de remarquer que :*

$$\{m : G_\gamma(m) \geq 0\} \supseteq \{m : A_\gamma(m) \geq 0\} \text{ pour tout } \gamma > 0.$$

Conjecture 3.2.1. (1) *Pour tout $\gamma > 0$, la condition nécessaire $G_\gamma(m) \geq 0$ se réalise si et seulement si $m \leq m_0(\gamma)$ où $m_0(\gamma)$ est la solution de l'équation $G_\gamma(m) = 0$ avec $G_\gamma(m)$ donné par le théorème précédent.*

(2) *Pour $\gamma > 0$, la loi a priori π_m est la moins favorable et l'estimateur de Bayes δ_m est minimax si et seulement si $m \leq m_0(\gamma)$.*

Remarque 3.2.2. *La partie* (2) *de la conjecture est obtenue en utilisant la partie* (1) *de la conjecture et le théorème 3.2.3.*

Nous vérifions que la partie (1) de la conjecture est vraie dans le cas particulier où $\gamma = 1$, par l'exemple suivant.

Exemple 3.2.4. *En utilisant la relation (3.2.12), pour $\gamma = 1$ et $m > 0$, nous avons :*

$$y(m) = \frac{\exp(-m) + m - 1}{1 - \exp(-m)}.$$

La fonction $G_1(m)$ est alors égale à :

$$\begin{aligned} G_1(m) &= 1-\exp(-m)-y(m), \\ &= \frac{\exp(-2m)-3\exp(-m)+2-m}{1-\exp(-m)}. \end{aligned}$$

Nous posons $k(m)=\exp(-2m)-3\exp(-m)+2-m$. Nous avons $k''(m) = \exp(-m)(4\exp(-m)-3)$. La fonction $k''(m)$ fait un changement de signe, elle passe du positif au négatif avec $k'(0)=0$ et $k(0)=0$. Ceci montre que $G_1(m)$ est positif si et seulement si $m \leq m_0(1)$, avec $m_0(1) \approx 1.15139$ est solution en m de l'équation $k(m)=0$.

Nous terminons cette section par la présentation des évaluations numériques de $G_\gamma(m)$, de la fonction de risque de δ_m, des fonctions $m_0(\gamma)$ et $m_1(\gamma)$ que nous accompagnons de diverses commentaires.

(1) Lorsque $\gamma > 0$, nous savons que la condition nécessaire et suffisante pour que la loi a priori soit concentrée sur $\{0, m\}$ est $G_\gamma(m) \geq 0$. Nous avons tenté en vain de démontrer que la condition $G_\gamma(m) \geq 0$ et $m \leq m_0(\gamma)$ sont équivalentes avec $m_0(\gamma)$ solution de l'équation de $G_\gamma(m)=0$. Nous avons alors procédé à l'évaluation numérique de l'ensemble $\{m : G_\gamma(m) \geq 0\}$ et nous constatons que cet ensemble coïncide avec l'intervalle $[0, m_0(\gamma)]$ (voir la figure 3.1). Nous savons aussi que la condition $A_\gamma(m) \geq 0$ est une condition nécessaire et suffisante pour que le risque de δ_m soit convexe, donc une condition suffisante pour que la loi a priori la moins favorable soit concentrée sur $\{0, m\}$. Nous avons évalué numériquement l'ensemble $\{m : A_\gamma(m) \geq 0\}$ et nous constatons que cet ensemble est égal à $[0, m_1(\gamma)]$ où $m_1(\gamma)$ est solution de l'équation $A_\gamma(m)=0$ et est représenté à la figure 3.4.

(2) Lorsque $\gamma < 0$, nous savons que la condition $m \leq m_1(\gamma)$ est une condition nécessaire et suffisante pour que le risque de δ_m soit convexe en θ, donc une condition suffisante pour que la loi a priori la moins favorable soit concentrée sur $\{0, m\}$. Mais, nous avons aussi évalué numériquement par la figure 3.4 les valeurs de $m_0(\gamma)$ pour lesquelles la loi a priori la moins

favorable soit concentrée sur $\{0, m\}$. Ces évaluations numériques sont basées sur l'algorithme suivant qui consiste à chercher les valeurs de m pour lesquelles π_m soit la moins favorable. Pour chaque valeur de γ fixée, nous procédons de la manière suivante :

(i) commencer avec un choix de m (par exemple $m = m_1(\gamma)$),

(ii) calculer la valeur de $y(m)$ correspondant en résolvant l'équation (3.2.12) pour la perte LINEX.

(iii) évaluer $\hat{\theta}$ tel que $\sup_{0 \leq \theta \leq m} R(\theta, \delta_m) = R(\hat{\theta}, \delta_m)$,

(iv) si $R(\hat{\theta}, \delta_m) \leq R(0, \delta_m)$ alors on choisit une autre valeur de m (par exemple $m + dl$ avec dl est le pas de selection, on le prend très petit) et on reprend le processus,

sinon, nous avons $m - dl \leq m_0(\gamma) < m$ et on arrête.

(3) Nous avons proposé quelques représentations de la fonction de risque de δ en fonction de θ pour différentes valeurs de m et γ dans des cas où δ_m est minimax et dans des cas où δ_m n'est pas minimax, par les figures 3.2 et 3.3 respectivement.

À partir de la figure 3.4 où sont représentées en fonction de γ les valeurs de $m_0(\gamma)$ et $m_1(\gamma)$, nous constatons que,

- $m_0(\gamma)$ croît en fonction de γ. C'est-à-dire que plus γ est grand, plus l'intervalle, pour lequel la loi a priori la moins favorable est concentrée sur $\{0, m\}$ est grand. Par exemple, si $\gamma = -3$, cet intervalle est donné par $[0, 0.570]$ et lorsque $\gamma = 3$, cet intervalle est $[0, 1.951]$. Remarquer que $m_0(\gamma) \approx 0.913$ lorsque $\gamma \approx 0$, ce qui rejoint le résultat pour la minimaxité pour la perte quadratique obtenu par Eichenauer (1986) et Berry (1993).

- lorsque γ est négatif, $m_1(\gamma)$ est proche de $m_0(\gamma)$ et croît en fonction de γ. Mais, lorsque γ est positif, au debut (pour de petites valeurs de γ), la différence entre $m_0(\gamma)$ et $m_1(\gamma)$ n'est pas trop grande, mais lorsque γ dévient grande, cette différence s'accentue. Par exemple, cette différence est égale à 0.226 lorsque $\gamma = 0.01$, elle est égale 1.044 lorsque $\gamma = 3$.

3.2.3. Distribution uniforme

Eichenauer et Fieger (1992) ont étudié le problème d'estimation minimax du paramètre θ, basé sur n observations d'une famille de distributions ayant la densité de la forme $f(x|\theta) = c(\theta)f_0(x)I_{[\theta,\theta+1]}(x)$ lorsque $0 \leq \theta \leq m$, m connu, pour une fonction de perte convexe. Ils montrent l'existence de m_0 tel que pour tout $m \leq m_0$, la loi a priori la moins favorable est concentrée sur $\{0, m\}$. Cependant, il ne donnent explicitement la valeur de m_0 que pour le cas d'une perte absolue et pour $f_0 \equiv 1$.

Dans cette partie, nous allons étudier le cas d'une distribution uniforme, $X \sim U(\theta, \theta + c)$ pour une perte quadratique et une perte LINEX. Nous allons proposer, pour les pertes quadratique et LINEX, des conditions nécessaires et suffisantes de la forme $m \leq m_0$ (en donnant des valeurs explicites de m_0) pour lesquelles la loi a priori la moins favorable est concenttrée sur $\{0, m\}$.

La fonction f_0 est alors donnée par :

$$f_0(x) = \begin{cases} c^{-1} & \text{si } 0 \leq x \leq c, \\ 0 & \text{sinon}, \end{cases}$$

où c est une constante. Nous supposons toujours que $c \geq m$ (sinon les estimateurs de Bayes ne seront pas uniques).

Pour une distribution uniforme, la loi a priori et la loi a posteriori sont données par les relations (3.2.4) et (3.2.11). Ainsi, k_α prend la forme suivante :

$$k_\alpha(x, m) = \begin{cases} \alpha & \text{si } m \leq x < c, \\ 0, & \text{si } c \leq x \leq m + c, \end{cases}$$

Il faut remarquer que sur les régions $[0, m)$, $[m, c)$ et $[c, m + c]$, cette perte a posteriori ne dépend pas de x. Ce qui simplifie la recherche d'estimateur de Bayes.

Pour tout $\alpha \in [0, 1]$, ceci procure à l'estimateur de Bayes $\delta_{\alpha,m}$ la forme suivante

$$\delta_{\alpha,m}(x) = \begin{cases} 0 & \text{si } 0 \leq x < m, \\ y_\alpha(m) & \text{si } m \leq x < c, \\ m & \text{si } c \leq x \leq m + c \end{cases} \tag{3.2.22}$$

avec $y_\alpha(m)$ minimisant l'expression $\rho(d)\alpha + \rho(d-m)(1-\alpha)$ en d.
Par la relation (3.2.10) et pour tout $\alpha \in [0,1]$, la fonction de risque de $\delta_{\alpha,m}$ devient :

$$R(\theta, \delta_{\alpha,m}) = c^{-1}[\rho(-\theta)(m-\theta) + \rho(y_\alpha(m)-\theta)(c-m) + \theta\rho(m-\theta)]. \quad (3.2.23)$$

Nous sommes à présent à la recherche de conditions pour lesquelles la loi a priori la moins favorable soit concentrée sur $\{0,m\}$. Mais, nous vérifions les conditions de l'appendice, donc une condition nécessaire pour laquelle la loi a priori π_α concentrée sur $\{0,m\}$ soit la moins favorable est :

$$R(0, \delta_{\alpha,m}) = R(m, \delta_{\alpha,m}). \quad (3.2.24)$$

En outre, si une telle condition s'est réalisée, nous dirons comme dans le cas exponentiel, que $\delta_{\alpha,m}$ est un "égalisateur". Un unique "égalisateur", pour chaque valeur de m donnée, est fourni par le lemme suivant.

Lemme 3.2.5. *Pour tous les estimateurs de la forme (3.2.22), il existe exactement un égalisateur donné par :*

$$\delta_m(x) = \begin{cases} 0 & si\ 0 \le x < m, \\ y(m) & si\ m \le x < c, \\ m & si\ c \le x \le m+c \end{cases}$$

où $y(m)$ satisfait à l'équation suivante en y :

$$\rho(y) = \rho(y-m). \quad (3.2.25)$$

DÉMONSTRATION. La démonstration est similaire à celle du lemme 3.2.2. □

Nous notons par π_m la loi a priori par rapport à l'estimateur de Bayes δ_m.

Exemple 3.2.5. *Pour tout $m \in [0,c]$, nous avons les résultats suivants.*

(1) *Pour une perte L^q, $q > 0$, la relation (3.2.25) est équivalente à*

$$y(m)^q = (m - y(m))^q.$$

La solution de l'équation précédente est égale à

$$y(m) = \frac{m}{2},$$

(observer que cette solution $y(m) = m/2$ sera celle pour une fonction de perte symétrique ρ).

(2) *Pour une perte LINEX, la relation (3.2.25) est équivalente à*

$$\exp(\gamma y(m)) - \gamma y(m) - 1 = \exp(\gamma(y(m) - m) - \gamma(y(m) - m) - 1.$$

L'unique solution de cette équation, pour tout $m \leq c$, est donnée par :

$$y(m) = \frac{1}{\gamma} \log\left(\frac{\gamma m}{1 - exp(-\gamma m)}\right). \tag{3.2.26}$$

Contrairement au cas exponentiel, nous avons pour le cas uniforme, une formule explicite de $y(m)$ pour la perte LINEX.

Nous proposons à présent, par les théorèmes 3.2.4 et 3.2.5 respectivement,

(1) une condition nécessaire et suffisante sur m donnée par $m \leq c/2$ pour que π_m soit la loi a priori la moins favorable et l'estimateur de Bayes δ_m soit minimax, pour la perte quadratique,

(2) des conditions nécessaires et suffisantes sur m de la forme $m \leq m_\gamma$ pour que π_m soit la loi a priori la moins favorable et δ_m soit minimaxe, pour une perte LINEX.

Dans le cas où la fonction de perte est quadratique et par la relation (3.2.23), la fonction de risque de δ_m devient :

$$\begin{aligned} R(\theta, \delta_m) &= c^{-1}[(c-2m)\theta^2 + (m^2 + 2my(m) - 2cy(m))\theta + (c-m)y^2(m)], \\ &= c^{-1}\left[(c-2m)\theta^2 + (2m^2 - cm)\theta + (c-m)\frac{m^2}{4}\right]. \end{aligned}$$

Cette fonction de risque est un polynôme de degré deux en θ. Il faut noter que trois possibilités peuvent se présenter pour le comportement de cette fonction de risque :

(1) si $m < c/2$ alors elle devient convexe,

(2) si $m > c/2$ alors elle devient concave,

(3) si $m = c/2$ elle devient constante.

Nous proposons par le théorème suivant les résultats de minimaxité.

Théorème 3.2.4. *Sous la fonction de perte quadratique, la loi a priori π_m est la moins favorable et l'estimateur de Bayes δ_m est minimax si et seulement si $m \leq c/2$.*

DÉMONSTRATION. Ce résultat découle du comportement de la fonction de risque qui est convexe si $m \leq c/2$ et concave sinon, du fait que $R(0, \delta_m) = R(m, \delta_m)$ et du théorème 3.3.8 (appendice). □

Étudions maintenant le cas où la fonction de perte est LINEX. Dans ce cas, par la relation (3.2.23), la fonction de risque de δ_m devient :

$$\begin{aligned} R(\theta, \delta_m) &= c^{-1}\{\exp(-\gamma\theta)[\theta(\exp(\gamma m) - 1) + m + (c-m)\exp(\gamma y)] \\ &\qquad -\gamma(c-m)(y-\theta) - c\}. \end{aligned} \tag{3.2.27}$$

Cette fonction de risque de δ_m et ses deux premières dérivées, nous permettent d'obtenir, par le prochain théorème, nos conditions necessaires et suffisantes dans le cas d'une perte LINEX.

Théorème 3.2.5. *Dans le cas où la fonction de perte est LINEX, nous avons les résultats suivants.*

(*a*) *Pour tout $\gamma \neq 0$, la fonction de risque de l'estimateur δ_m est :*

- *convexe si $m \leq m_\gamma$,*
- *concave si $m > m_\gamma$,*

avec m_γ solution de l'équation $H_\gamma(m) = 0$ où

$$H_\gamma(m) = |\gamma|\{-2(\exp(|\gamma|m) + \exp(-|\gamma|m) - 2) + |\gamma|m(1 - \exp(-|\gamma|m)) + |\gamma|^2 m(c-m)\}.$$

(*b*) *Pour tout $\gamma \neq 0$, l'estimateur de Bayes δ_m est minimax et la loi a priori π_m est la moins favorable si et seulement si $m \leq m_\gamma$.*

DÉMONSTRATION. (*a*) Par la relation (3.2.27), la dérivée deuxième de la fonction de risque de l'estimateur δ_m est égale à :

$$\begin{aligned} \frac{\partial^2}{\partial\theta^2} R(\theta, \delta_m) &= c^{-1}\exp(-\gamma\theta)\{\gamma^2\theta(\exp(\gamma m) - 1) + \gamma^2(m + (c-m)\exp(\gamma y)) \\ &\qquad -2\gamma(\exp(\gamma m) - 1)\}, \\ &= \exp(-\gamma\theta) Q_\gamma(\theta, m), \end{aligned}$$

avec

$$Q_\gamma(\theta, m) \quad = \quad \gamma^2\theta(\exp(\gamma m) - 1) + \gamma^2(m + (c - m)\exp(\gamma y(m))) - 2\gamma(\exp(\gamma m) - 1).$$

En remplaçant les valeurs de $y(m)$ (relation 3.2.26) dans cette dernière égalité, nous avons :

$$Q_\gamma(\theta, m) \quad = \quad \frac{(\gamma^2\theta - 2\gamma)(\exp(\gamma m) + \exp(-\gamma m) - 2) + \gamma^2 m(1 - \exp(-\gamma m)) + \gamma^3 m(c - m)}{1 - \exp(-\gamma m)}.$$

Il est facile de voir que la fonction $Q_\gamma(\theta, m)$ est croissante en θ si $\gamma > 0$ et décroissante en θ si $\gamma < 0$. Nous avons alors :

$$\frac{\partial^2}{\partial\theta^2} R(\theta, \delta_m) \quad \geq \quad \begin{cases} \exp(-\gamma\theta) Q_\gamma(0, m) \text{ si } \gamma > 0, \\ \exp(-\gamma\theta) Q_\gamma(m, m) \text{ si } \gamma < 0. \end{cases}$$

et

$$\frac{\partial^2}{\partial\theta^2} R(\theta, \delta_m) \quad \leq \quad \begin{cases} \exp(-\gamma\theta) Q_\gamma(m, m) \text{ si } \gamma > 0, \\ \exp(-\gamma\theta) Q_\gamma(0, m) \text{ si } \gamma < 0. \end{cases}$$

Cependant, nous pouvons écrire :

$$H_\gamma(m) = \begin{cases} Q_\gamma(0, m)[1 - \exp(-\gamma m)] \text{ si } \gamma > 0 \\ Q_\gamma(m, m)[\exp(-\gamma m) - 1] \text{ si } \gamma < 0. \end{cases}$$

En outre, il est facile de voir que $H_\gamma(m) = H_{-\gamma}(m)$, pour tout $m > 0$ et $\gamma \neq 0$. Alors, $Q_\gamma(\theta)$ est positif (respectivement négatif) pour tout θ si et seulement si $H_\gamma(m)$ est positif (respectivement négatif).

Mais, il faut noter que cette fonction est concave avec $H_\gamma(0) = 0$. En effet, nous avons, pour tout $m \geq 0$,

$$\begin{aligned} \frac{\partial^2}{\partial m^2} H_\gamma(m) \quad &= \quad |\gamma|\{-|\gamma|^3 m \exp(-|\gamma| m) - 2|\gamma|^2(1 + \exp(|\gamma| m))\} \\ &\leq \quad 0. \end{aligned}$$

Donc, cette fonction fait un changement de signe en passant du positif au négatif, ce qui nous donne le résultat.

(b) C'est une conséquence de la partie (a), du fait que $R(0, \delta_m) = R(m, \delta_m)$ et du théorème 3.3.8 (appendice). □

Remarque 3.2.3. (*a*) *Quelques valeurs de m_γ en fonction de γ, pour $c = 1$, sont données par la figure 3.5. Lorsque m est assez petit, la loi a priori la moins favorable est concentrée sur $\{0, m\}$. Nous constatons que la fonction m_γ est symétrique en γ autour de 0. Ceci apparaît dans la démonstration des résultats (par la fonction $H_\gamma(m)$). De plus, elle décroît en fonction de $|\gamma|$ et atteint son maximum à $\gamma = 0$. Il faut aussi noter que $\lim_{\gamma \to 0} m_\gamma = 1/2$, c'est-à-dire la constante obtenue pour le cas de la perte L^2.*

(*b*) *Il faut noter que nous avons traité le cas uniforme que pour un échantillon de taille $n = 1$. La généralisation des résultats obtenus dans le cas d'un échantillon de taille $n > 1$ sera sûrement très compliquée.*

3.3. Distribution discrète

Dans cette section, nous traitons le problème d'estimation minimax d'un paramètre $\theta \in \Theta(m) = [0, m]$, $m > 0$ basé sur des observations issues d'une famille de distributions discrètes de fonction de masse p_θ. Nous commençons par proposer des propriétés générales sous une fonction de perte strictement convexe. Ensuite, dans le cas particulier d'une fonction de perte quadratique et sous certaines conditions sur la distribution, nous proposons des conditions sur m pour que la loi a priori la moins favorable soit concentrée sur $\{0, m\}$. Enfin, dans le cas d'une perte LINEX et pour une distribution de Poisson, nous proposons des conditions sur m pour lesquelles la loi a priori la moins favorable soit concentrée sur $\{0, m\}$. Dans les deux cas, nous proposons soit des conditions nécessaires et suffisantes, soit des conditions seulement suffisantes. Nous reprenons les résultats de Marchand et Parsian (2002, manuscrit non publié) pour les propriétés générales et le premier cas d'une perte quadratique. Les résultats obtenus pour la perte quadratique serons appliqués aux cas particuliers d'une distribution de Poisson, d'une distribution Binomiale et d'une distribution Binomiale négative.

3.3.1. Préliminaires, notations et propriétés générales

Soient $X_1, X_2, \ldots, X_n$ des variables aléatoires indépendantes identiquement distribuées de fonction de masse p_θ, $\theta \geq 0$ telles que :

(i) $P_\theta(X_i = 0) = p_\theta(0)$ est décroissant en θ avec $p_0(0) = 1$,

(ii) le support de X_i est soit égal à $\{0, 1, 2, ..., s\}$ ou soit égal à $\{0, 1, 2, ...\}$.

Ces distributions sont assez générales. Par exemple, si nous considérons la famille des distributions dite " Modified Power Series" (voir Johnson, Kotz et Kemp (1992)). La fonction de masse d'une distribution d'une telle famille est de la forme :

$$p_\theta(x) = c(\theta)h(x)[g(\theta)]^x,$$

avec $c(\theta) = \left[\sum_x h(x)[g(\theta)]^x\right]^{-1}$. Mais, la fonction $p_\theta(0) = c(\theta)h(0)$ est décroissante en θ si g est croissant en θ. Nous posons $c(0) = [h(0)]^{-1}$. Nous avons alors $p_0(0) = 1$. Dans cette grande famille de distributions, on retrouve, parmi d'autres, les distributions suivantes :

- Binomiale (n, θ) : $c(\theta) = (1-\theta)^n$, $g(\theta) = \theta/(1-\theta)$ et $h(x) = \binom{n}{x}$.
- Poisson (θ) : $c(\theta) = \exp(-\theta)$, $g(\theta) = \theta$ et $h(x) = 1/x!$.
- Binomiale négative (n, θ) : la fonction de masse est $P(X = x|p) = \binom{n+x-1}{x} p^n (1-p)^x I_{I\!N}(x)$. On pose $1 - p = \theta/(n+\theta)$. Nous avons $c(\theta) = (n/(n+\theta))^n$, $g(\theta) = \theta/(n+\theta)$ et $h(x) = \binom{n+x-1}{x}$. Pour $n = 1$, on retrouve la distribution Géométrique.

Soit la fonction de perte strictement convexe donnée par l'expression (3.2.2). Pour $\alpha \in [0, 1]$, on considère la loi a priori concentrée sur $\{0, m\}$ donnée par :

$$\pi_\alpha(0) = 1 - \pi_\alpha(m) = \alpha.$$

La loi a posteriori est également concentrée sur la frontière $\{0, m\}$ et elle est donnée par :

$$\pi_\alpha(0|x) = 1 - \pi_\alpha(m|x) = \begin{cases} \psi_\alpha(m) & \text{si } \sum_{i=1}^n x_i = 0, \\ 0 & \text{sinon.} \end{cases}$$

avec $x = (x_1, x_2, ..., x_n)$ et $\psi_\alpha(m) = \alpha/(\alpha + (1-\alpha)p_m^n(0))$.

Ainsi, la perte a posteriori est égale à :

$$E[\rho(d-\theta)|x] = \begin{cases} \psi_\alpha(m)\rho(d) + (1-\psi_\alpha(m))\rho(d-m) & \text{si } \sum_{i=1}^n x_i = 0, \\ \rho(d-m) & \text{sinon.} \end{cases}$$

Cette perte a posteriori est minimisée par le choix de $d = m$ et $d = y_\alpha(m)$ suivant les valeurs de x où $y_\alpha(m)$ est l'unique valeur qui minimise en d l'expression suivante :

$$\psi_\alpha(m)\rho(d) + (1 - \psi_\alpha(m))\rho(d - m). \tag{3.3.1}$$

En conséquence, la loi a priori concentrée sur les points $\{0, m\}$ donne un estimateur de Bayes de la forme :

$$\delta_{\alpha,m}(x) = \begin{cases} y_\alpha(m) & \text{si } \sum_{i=1}^{n} x_i = 0, \\ m & \text{sinon .} \end{cases} \tag{3.3.2}$$

Nous proposons, par l'exemple suivant, les estimateurs de Bayes $\delta_{\alpha,m}$ dans les cas particuliers d'une perte L^q, $q > 1$ et d'une perte LINEX.

Exemple 3.3.1. *(1) Pour la perte L^q avec $q > 1$, nous avons :*

$$y_\alpha(m) = m\left[1 + \left(\frac{\alpha}{1-\alpha} p_m^{-n}(0)\right)^{1/(q-1)}\right]^{-1}. \tag{3.3.3}$$

(2) Pour la perte LINEX, nous avons :

$$y_\alpha(m) = \frac{1}{\gamma}\log\left(\frac{\alpha + (1-\alpha)p_m^n(0)}{\alpha + (1-\alpha)p_m^n(0)\exp(-\gamma m)}\right). \tag{3.3.4}$$

Les résultats de l'exemple précédent ont été obtenus en calculant, pour $\rho(x) = |x|^q$ et $\rho(x) = \exp(\gamma x) - \gamma x - 1$, la dérivée première de l'expression (3.3.1) en d et en cherchant la valeur de d pour laquelle cette dérivée première est nulle. Par la convexité de ρ, nous avons bien un minimum.

Nous mesurons la performance de l'estimateur de Bayes $\delta_{\alpha,m}$ par la fonction de risque de cet estimateur et elle est donnée par :

$$R(\theta, \delta_{\alpha,m}) = \rho(m - \theta) + p_\theta^n(0)[\rho(y_\alpha(m) - \theta) - \rho(m - \theta)]. \tag{3.3.5}$$

Nous somme présentement à la recherche de conditions pour lesquelles la loi a priori la moins favorable soit concentrée sur $\{0, m\}$. Mais nous savons, par le théorème 3.3.5 et le lemme 3.3.4 (appendice), qu'il est nécessaire que la fonction de risque de l'estimateur de Bayes $\delta_{\alpha,m}$ évaluée en $\theta = 0$ soit égale à celle évaluée en $\theta = m$ (c'est-à-dire $R(0, \delta_{\alpha,m}) = R(m, \delta_{\alpha,m})$) pour qu'une loi a priori π_α concentrée sur les deux points $\{0, m\}$ soit la moins favorable. Il s'en suit que si un estimateur $\delta_{\alpha,m}$ satisfait à cette condition, nous l'appellerons "égalisateur". Le prochain lemme nous procure un unique "égalisateur", pour chaque valeur de $m > 0$.

Lemme 3.3.1. *Pour tous les estimateurs de la forme (3.3.2) et pour tout $m > 0$, il existe exactement un "égalisateur" donné par :*

$$\delta_m(x) = \begin{cases} y(m) & si \sum_{i=1}^{n} x_i = 0, \\ m \ sinon , & \end{cases}$$

où $y(m)$ satisfait à l'équation suivante en y :

$$\rho(y) = \rho(y - m)p_m^n(0). \tag{3.3.6}$$

Démonstration. En utilisant la relation (3.3.5) pour le risque de $\delta_{\alpha,m}$, nous avons :

$$R(0, \delta_{\alpha,m}) = \rho(y_\alpha(m)) \text{ et } R(m, \delta_{\alpha,m}) = \rho(y_\alpha(m) - m)p_m^n(0).$$

La relation (3.2.11) est alors équivalente à la relation suivante :

$$\rho(y_\alpha(m)) = \rho(y_\alpha(m) - m)p_m^n(0).$$

Puisque $y_{\alpha,m}$ prend toutes les valeurs de $[0, m]$ lorsque α varie dans $[0, 1]$, la relation précédente est équivalente à la relation (3.3.6).
Soit la fonction $\Delta(y) = \rho(y) - \rho(y - m)p_m^n(0)$, pour $y \in [0, m]$. Cette fonction est strictement croissante sur $[0, m]$ avec $\Delta(0) = -\rho(-m)p_m^n(0) \leq 0$ et $\Delta(m) = \rho(m) \geq 0$. Il existe alors un unique $y \in [0, m]$ tel que $\Delta(y) = 0$. □

Nous présentons, par l'exemple suivant, les formes des "égalisateurs" pour la perte L^q, $q > 1$.

Exemple 3.3.2. *Pour la perte L^q, $q \geq 1$, l'expression (3.3.6) devient*

$$y^q(m) = (m - y(m))^q p_m^m(0).$$

En divisant des deux côtés de l'expression précédente par $y^q(m)$, nous pouvons tirer la valeur de $y(m)$. Elle est donnée par :

$$y(m) = m\frac{p_m^{n/q}(0)}{1+p_m^{n/q}(0)}. \tag{3.3.7}$$

Quoiqu'il en n'est pas nécessaire, nous pouvons déduire de cette relation et de la partie (1) de l'exemple 3.3.1, la valeur de α correspondante à $y(m)$. Elle est égale à :

$$\alpha_m = \frac{p_m^{n/q}(0)}{1+p_m^{n/q}(0)} = \frac{y(m)}{m}, \quad m > 0.$$

Le lemme qui suit, permet d'avoir quelques propriétés de la fonction y qui nous serons utiles par la suite.

Lemme 3.3.2. *Pour tout $n \geq 1$, s'il existe $m_0(n)$ telle que, pour tout $m \leq m_0(n)$, la loi a priori concentrée sur $\{0, m\}$ est la moins favorable alors*

(1) *la fonction $y(m)$ est croissante en m,*

(2) *$y'(m) \leq 1$.*

Démonstration. Nous savons que si la loi a priori π_m est la moins favorable, alors nous aurons les égalités suivantes pour le risque minimax $\overline{R}(m)$:

$$\overline{R}(m) = R(0, \delta_m) = R(m, \delta_m). \tag{3.3.8}$$

La relation précédente est équivalente à :

$$\overline{R}(m) = \rho(y(m)) = \rho(y(m) - m)p_m^n(0). \tag{3.3.9}$$

(1) D'une part, pour $m \leq m_0(n)$, la dérivée du risque minimax de δ_m est égale à :

$$\frac{\partial}{\partial m}\overline{R}(m) = y'(m)\rho'(y(m)).$$

Par le lemme 3.2.1, $y'(m)\rho'(y(m))$ est positif. Comme la fonction de perte ρ est strictement convexe (donc $\rho'(x) > 0$ pour $x > 0$), alors nous avons $y'(m) \geq 0$.

(2) D'autre part, pour $m \leq m_0(n)$, la dérivée du risque minimax est égale à

$$\frac{\partial}{\partial m}\overline{R}(m) = (y'(m) - 1)\rho'(y(m) - m)p_m^n(0) + \rho(y(m) - m)[p_m^n(0)]'.$$

Par le lemme 3.2.1, $(y'(m)-1)\rho'(y(m)-m)p_m^n(0)+\rho(y(m)-m)[p_m^n(0)]'$ est positif. Mais, comme la fonction ρ est strictement convexe alors $\rho'(x)<0$ pour $x<0$ et $\rho(x)>\rho(0)=0$. En utilisant alors le fait que $p_m^n(0)$ est décroissant en m, nous avons $y'(m)-1\leq 0$. Ce qui achève la démonstration. $\square$

Nous allons proposer, par le théorème suivant, des conditions nécessaires pour que

$$\sup_{0\leq\theta\leq m} R(\theta,\delta_m)=\max\{R(0,\delta_m),R(m,\delta_m)\}=R(0,\delta_m)=R(m,\delta_m). \quad (3.3.10)$$

Théorème 3.3.1. *Chacune des deux conditions suivantes sont nécessaires pour que la condition (3.3.10) se réalise (c'est-à-dire que δ_m est minimax).*

(1)

$$\rho(m)-\rho(y(m))\leq\frac{c}{n}\rho'(y(m)), \quad (3.3.11)$$

avec $c=-\left[\left(\frac{\partial}{\partial\theta}p_\theta(0)\big|_{\theta=0}\right)\right]^{-1}$.

(2)

$$\rho(y(m)-m)\leq\frac{d}{n}\rho'(y(m)-m), \quad (3.3.12)$$

avec $d=p_m(0)/\left[\left(\frac{\partial}{\partial\theta}p_\theta(0)\big|_{\theta=m}\right)\right]=\left[\left(\frac{\partial}{\partial\theta}\log p_\theta(0)\ \big|_{\theta=m}\right)\right]^{-1}$.

Démonstration. Il est évident que si la condition (3.3.10) est vraie alors la fonction de risque de δ_m doit être décroissante en $\theta=0$, et doit être croissante en $\theta=m$. Mais, en utilisant l'expression (3.3.5) pour le risque de δ_m, nous avons :

$$\frac{\partial}{\partial\theta}R(\theta,\delta_m)\bigg|_{\theta=0} = \frac{n}{c}[\rho(m)-\rho(y(m))]-\rho'(y(m)),$$

et

$$\frac{\partial}{\partial\theta}R(\theta,\delta_m)\bigg|_{\theta=m} = p_m^n(0)\left\{\frac{n}{d}\rho(y(m)-m)-\rho'(y(m)-m)\right\}.$$

Ce qui achève la démonstration. $\square$

Nous appliquons les résultats du théorème précédent aux cas d'une fonction de perte L^q, $q>1$ et d'une fonction de perte LINEX. Les résultats obtenus sont donnés par les deux prochains corollaires respectivement.

Corollaire 3.3.1. *Pour une perte L^q, $q > 1$ et pour $n \geq 1$, la condition nécessaire (3.3.11) est réalisée si et seulement si $m \leq m_1(n)$, avec $m_1(n)$ solution de l'équation en m suivante :*

$$m - \frac{cq}{n}\frac{b(m)}{b^q(m) - 1} = 0, \text{ avec } b(m) = \frac{1 + p_m^{n/q}(0)}{p_m^{n/q}(0)} \tag{3.3.13}$$

Ceci veut dire que lorsque $m > m_1(n)$, alors l'estimateur de Bayes δ_m n'est pas minimax.

Démonstration. Pour une perte L^q, la condition nécessaire (3.3.11) devient :

$$m^q - y^q(m) \leq \frac{cq}{n} y^{q-1}(m). \tag{3.3.14}$$

En divisant la relation précédente par m^q (avec $m \neq 0$), nous avons :

$$m \leq \frac{cq}{n} \frac{\left(\frac{y(m)}{m}\right)^{q-1}}{1 - \left(\frac{y(m)}{m}\right)^q}.$$

Considérons la fonction suivante

$$h(m) = m - \frac{cq}{n} \frac{\left(\frac{y(m)}{m}\right)^{q-1}}{1 - \left(\frac{y(m)}{m}\right)^q}. \tag{3.3.15}$$

Il est facile de voir, par la relation (3.3.7), que la fonction $y(m)/m$ est décroissante en m, donc $h(m)$ est croissante en m avec $h(0) = -2cq/(n(2^q - 1)) \leq 0$. Ce qui achève la démonstration. □

Remarque 3.3.1. *Pour la fonction de perte quadratique, nous avons :*

$$h(m) = \frac{mn - 2(mn - c)p_m^{n/2}(0) - 2cp_m^n(0)}{n(1 + 2p_m^{n/2}(0))}.$$

Par conséquent, l'équation (3.3.13) est équivalente à l'équation suivante :

$$mn + 2[mn - c]p_m^{n/2}(0) - 2cp_m^n(0) = 0, \text{ avec } c = \left[\frac{\partial}{\partial\theta} p_\theta(0)\bigg|_{\theta=0}\right]^{-1}. \tag{3.3.16}$$

Corollaire 3.3.2. *Pour une perte LINEX et une distribution de Poisson, nous avons les résultats suivants :*

(1) *si $\gamma \leq -n$, la condition nécessaire (3.3.12) se réalise pour tout $m > 0$,*

(2) *si $\gamma > -n$, cette condition se réalise si et seulement si $m \leq m^*(n, \gamma)$, où*

$m^*(n,\gamma)$ solution de l'équation en $h_\gamma(m) = 0$
où $h_\gamma(m) = (n+\gamma)(\exp(\gamma(y(m)-m)) - 1) - n\gamma(y(m)-m)$.

DÉMONSTRATION. Pour tout $n \geq 1$, la condition (3.3.12) devient :

$$(n+\gamma)(\exp(\gamma(y(m)-m)) - 1) - n\gamma(y(m)-m) \leq 0. \tag{3.3.17}$$

Cette relation est équivalente à $h_\gamma(m) \leq 0$.

- Si $\gamma \leq -n$, pour tout $n \geq 1$, il est facile de voir que $h_\gamma(m) \leq 0$, pour tout $m \geq 0$.
- Nous supposons que $\gamma > -n$. La dérivée de la fonction h_γ est donnée par $h'_\gamma(m) = \gamma(y'(m)-1)[(n+\gamma)\exp(\gamma(y(m)-m)) - n]$, pour $m \geq 0$. Par le lemme 3.3.2, nous avons $y'(m) - 1 \leq 0$ donc la fonction $h_\gamma(m)$ et $k_\gamma(m) = -\gamma[(n+\gamma)\exp(\gamma(y(m)-m)) - n]$ ont le même signe. Mais, la fonction $k_\gamma(m)$ est croissante en m avec $k_\gamma(0) = -\gamma^2 \leq 0$. La fonction $h_\gamma(m)$ passe alors du négatif au positif. Mais, il est facile de voir avec la relation (3.3.4) que $\lim_{m\to\infty} y(m) = +\infty$. Il s'en suit alors que $\lim_{m\to\infty} h_\gamma(m) > 0$. En utilisant le fait que $h_\gamma(0) = 0$, nous avons $h_\gamma(m)$ qui passe du négatif au positif. Ceci nous donne le résultat.

□

Nous allons à présent proposer, à la sous-section suivante, par le théorème 3.3.2, pour la fonction de perte quadratique et sous certaines conditions sur la distribution :

(A) soit des conditions nécessaires et suffisantes sur m pour que π_m soit la loi a priori la moins favorable et δ_m soit minimax.

(B) soit des conditions seulement suffisantes sur m pour que π_m soit la loi a priori la moins favorable et δ_m soit minimax.

Nos conditions nécessaires et suffisantes et nos conditions suffisantes sont obtenus en montrant que la deuxième dérivée par rapport à θ de la fonction de risque de δ_m est croissante, pour tout $m \leq m_2(n)$ avec $m_2(n)$ défini par le théorème 3.3.2. Ce comportement de la deuxième dérivée conjugué à la contrainte $R(0,\delta_m) = R(m,\delta_m)$ nous permet de travailler avec la condition $\frac{\partial}{\partial\theta}R(\theta,\delta_m)\big|_{\theta=0} \leq 0$. Mais, nous avons montré par le corollaire 3.3.1, que cette dernière condition est équivalente à $m \leq m_1(n)$, pour tout $n \geq 1$. Alors, la condition $m \leq m_1(n) \wedge m_2(n)$ devient une condition suffisante. Et, de plus lorsque $m_2(n) \geq m_1(n)$, cette dernière condition devient $m \leq m_1(n)$, c'est-à-dire une condition nécessaire et suffisante. Nous appliquons ces résultats ensuite aux cas d'une distribution de Poisson, d'une distribution de Binomiale et d'une distribution Binomiale négative.

3.3.1.1. *Perte quadratique*

Dans cette sous-section, nous supposons en plus des conditions (i) et (ii), les conditions suivantes :

(iii) la fonction $p_\theta(0)$ derivable en θ et de dérivée continue en θ, $\frac{\partial}{\partial\theta}p_\theta(0)$ derivable en θ et de dérivée continue,

(iv) $\frac{\partial}{\partial\theta}p_\theta^n(0)$ croissante en θ,

(v) $\frac{\partial^2}{\partial\theta^2}p_\theta^n(0)$ décroissante en θ.

Nous posons $z(\theta) = \theta\frac{\partial^2}{\partial\theta^2}p_\theta^n(0)$ et $m_2(n) = \inf\{\theta : z'(\theta) < 0\}$. Il faut noter que $m_2(n) > 0$ puisque $z'(0) = \frac{\partial^2}{\partial\theta^2}p_\theta^n(0)\Big|_{\theta=0} + \lim_{\theta\to 0}\theta\frac{\partial^3}{\partial\theta^3}p_\theta^n(0) = \frac{\partial^2}{\partial\theta^2}p_\theta^n(0)\Big|_{\theta=0} \geq 0$, pour tout $n \geq 1$.

Ces conditions de régularité sur p_θ et ses dérivées (conditions (i) à (v)) nous servent à démontrer, par le prochain théorème, que la troisième dérivée de la fonction de risque de δ_m est positive pour m suffisamment petit.

Théorème 3.3.2. *Sous la fonction de perte quadratique, nous avons les résultats suivants :*

(1) Pour tout $\theta \in \Theta(m)$ et $m \leq m_2(n)$, la dérivée troisième de la fonction de risque est positive, c'est-à-dire, pour tout $m \leq m_2(n)$,

$$\frac{\partial^3}{\partial\theta^3}R(\theta, \delta_m) \geq 0, \quad \forall\theta \in \Theta(m).$$

(2) Si $m \leq m_0(n) = m_1(n) \wedge m_2(n)$ alors la loi a priori π_m est la moins favorable et l'estimateur de Bayes δ_m est minimax.

(3) Si $m_1(n) \leq m_2(n)$, la condition $m \leq m_1(n)$ est nécessaire et suffisante pour que la loi a priori π_m soit la moins favorable et l'estimateur de Bayes δ_m soit minimax.

Démonstration. (1) Pour une perte quadratique, la fonction de risque de δ_m devient :

$$R(\theta, \delta_m) = (m-\theta)^2 + p_\theta^n(0)[m - y(m)][2\theta - m - y(m)].$$

Les deux premières dérivées du risque de δ_m sont données par :

$$\frac{\partial}{\partial\theta} R(\theta, \delta_m) = -2(m-\theta) + \frac{\partial}{\partial\theta} p_\theta^n(0)[m - y(m)][2\theta - m - y(m)] + 2p_\theta^n(0)[m - y(m)]$$

et

$$\frac{\partial^2}{\partial\theta^2} R(\theta, \delta_m) = 2 + [m - y(m)]\left[4\frac{\partial}{\partial\theta} p_\theta^n(0) + 2z(\theta) - (m + y(m))\frac{\partial^2}{\partial\theta^2} p_\theta^n(0)\right].$$

En utilisant les hypothèses de cette sous-section, nous avons la deuxième dérivée croissante en θ pour tout $m \leq m_2(n)$.

(2) C'est une conséquence du résultat (1) et de la discussion du dernier paragraphe à la fin de la section précédente.

(3) C'est une conséquence du résultat (2), de la discussion du dernier paragraphe à la fin de la section précédente et du théorème 3.3.8 (appendice).

□

Nous allons à présent appliquer ces résultats aux cas d'une distribution d'une distribution de Poisson, d'une distribution Binomiale et d'une distribution Binomiale négative.

Exemple 3.3.3. *: Distribution de Poisson.*

La fonction de masse d'une distribution de Poisson est donnée par :

$$p_\theta(x) = \frac{\exp(-\theta)\theta^x}{x!}, \quad x = 0, 1, \ldots \tag{3.3.18}$$

Nous avons alors $p_\theta^n(0) = \exp(-n\theta)$, $\frac{\partial}{\partial\theta} p_\theta^n(0) = -n\exp(-n\theta)$, $\frac{\partial^2}{\partial\theta^2} p_\theta^n(0) = n^2\exp(-n\theta)$. *Les conditions (i) à (v) sont alors toutes vérifiées. Ainsi, nous pouvons appliquer le théorème précédent pour obtenir les résultats de minimaxité. Ces résultats sont donnés par le corollaire suivant.*

Corollaire 3.3.3. *Pour une distribution de Poisson, et pour une perte quadratique, l'estimateur de Bayes δ_m est minimax et π_m est la loi a priori la moins favorable si et seulement si $m \leq m_0(n)$ avec $m_0(n) = y_0/n$ et $y_0 \approx 0.912955$ est solution en y de l'équation $y\exp(y) + 2(y-1)\exp(y/2) - 2 = 0$.*

DÉMONSTRATION. Nous avons $z'(\theta) = n^2 \exp(-n\theta)[1 - n\theta]$. Ce qui nous donne $m_2(n) = 1/n$. Par le théorème 3.3.2, l'estimateur de Bayes δ_m est minimax pour $m \leq m_0(n) = m_1(n) \wedge m_2(n)$, pour tout $n \geq 1$. Mais, nous avons montré, par la démonstration du corollaire (3.3.1), que $h(m) \leq 0$ si et seulement si $m \leq m_1(n)$. Comme $h(m_2(n)) = (e-2)/(n(e + 2e^{1/2})) > 0$, pour tout $n \geq 1$, alors nous avons $m_1(n) \leq m_2(n)$. Alors, par la partie (3) du théorème 3.3.2, la condition $m \leq m_1(n)$ est nécessaire et suffisante pour que π_m soit la moins favorable. Par la remarque 3.3.1, l'équation $h(m) = 0$ est équivalente à
$mn \exp(mn) + 2(mn-1)\exp(mn/2) - 2 = 0$. Si on pose $y = mn$ dans l'équation précédente, nous avons le résultat. □

Exemple 3.3.4. *: Distribution Binomiale*
La fonction de masse d'une distribution de Bernoulli est donnée par :

$$p_\theta(x) = \theta^x(1-\theta)^{1-x}, \quad x = 0, 1.$$

Nous avons alors $p_\theta^n(0) = (1-\theta)^n$, $\frac{\partial}{\partial\theta} p_\theta^n(0) = -n(1-\theta)^{n-1}$,
$\frac{\partial^2}{\partial\theta^2} p_\theta^n(0) = n(n-1)(1-\theta)^{n-2}$. *Les conditions (i) à (v) sont toutes vérifiées, nous utilisons alors le théorème précédent obtenir nos résultats de minimaxité. Ces résultats sont donnés par le corollaire suivant.*

Corollaire 3.3.4. *Pour une distribution Bernoulli et pour une perte quadratique, l'estimateur de Bayes δ_m est minimax et π_m est la loi a priori la moins favorable si et seulement si $m \leq m_0(n)$ avec $m_0(n)$ solution de l'équation*
$mn + 2(mn-1)(1-m)^{n/2} - 2(1-m)^n = 0$, *pour* $n \geq 1$.

DÉMONSTRATION. Nous avons pour $n = 1$, $z'(\theta) = 0$ et pour $n > 1$,
$z'(\theta) = n(n-1)(1-\theta)^{n-3}[1-(n-1)\theta]$. Ceci nous donne $m_2(n) = \infty$, si $n = 1$ et $m_2(n) = 1/(n-1)$, pour $n > 1$. Par le théorème 3.3.2, l'estimateur de Bayes δ_m est minimax pour $m \leq m_0(n) = m_1(n) \wedge m_2(n)$. Mais, nous savons, pour tout $n \geq 1$, que $h(m) \leq 0$ si et seulement si $m \leq m_1(n)$.
Comme $h(1/n) = (1 - 2(1-1/n)^n)/(n(1 + 2(1-1/n)n/2)) > 0$, pour tout $n \geq 1$ alors nous avons $m_1(n) \leq 1/n < m_2(n)$. Mais, par la remarque 3.3.1, l'équation $h(m) = 0$ est équivalente à $mn + 2(mn-1)(1-m)^{n/2} - 2(1-m)^n = 0$. Par le théorème 3.3.2 partie (3), nous avons le résultat. □

Remarque 3.3.2. *Il faut remarquer que le corollaire précédent englobe aussi le cas $n = 1$, mais d'une autre façon plus directe, nous avons* $\frac{\partial^2}{\partial\theta^2}R(\theta, \delta_m) = 2 - 4(m - y(m))$. *Ce qui montre que la condition* $m \leq m_0(1)$, *avec* $m_0(1) \approx 0.750$ *solution de l'équation* $1 - 2m + (1 - m)^{1/2} = 0$, *est une condition nécessaire et suffisante pour que* π_m *soit la loi a priori la moins favorable et* δ_m *soit minimax dans le cas où* $n = 1$.

Remarque 3.3.3. *Nous proposons, par le tableau suivant quelques valeurs de* $m_0(n)$. *Nous observons que l'espace des paramètres* $\Theta(m)$ *doit être assez petit pour que la loi a priori la moins favorable soit concentrée sur la frontière de* $\Theta(m)$. *De plus, nous avons* $m_0(n) \leq 1/n$ *et* $m_0(n)$ *est décroissante en* n *(ceci vient du fait que* $n(1 + 2(1 - m)^{n/2})h(m)$ *est croissante en* n*). Nous posons, pour* $n \geq 1$, $y_n = nm_0(n)$ *et* $\lim_{n\to\infty} y_n = y_*$. *Nous pouvons alors vérifier que* $\lim_{n\to\infty}(y_n + 2(y_n - 1)(1 - y_n/n)^{n/2} - 2(1 - y_n/n)^n) = 0$. *Alors,* y_* *serait solution de l'équation en* y *de* $y\exp(y) + 2(y - 1)\exp(y/2) - 2 = 0$. *La valeur de* y_* *est approximativement égale à* 0.912955. *Lorsque* n *est assez grand, nous avons* $m_0(n) \approx 0.912955/n$.

Il faut signaler que cet exemple a été traité par Marchand et MacGibbon (2000). Mais, nous avons traité cet exemple à partir d'un cadre plus général. Ces derniers ont aussi proposer un graphe qui donnent des valeurs de $m_0(n)$.

TAB. 3.1. Valeurs de $m_0(n)$ pour une distribution Bernoulli avec la perte quadratique.

n	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
$m_0(n)$	*0.750*	*0.423*	*0.290*	*0.220*	*0.178*	*0.149*	*0.128*	*0.112*	*0.100*	*0.090*
n	*15*	*20*	*30*	*40*	*50*	*60*	*70*	*80*	*90*	*100*
$m_0(n)$	*0.060*	*0.045*	*0.030*	*0.023*	*0.018*	*0.015*	*0.013*	*0.011*	*0.010*	*0.009*

Exemple 3.3.5. *: Distribution Binomiale négative*

La fonction de masse d'une distribution de géométrique est :

$P(X = x|p) = p(1 - p)^x, \quad x \in \mathbb{N}^*, p \in (0, 1]$ *avec* $E_p[X_i] = (1 - p)/p$.

On veut estimer le paramètre $\theta = (1-p)/p$. La fonction de masse devient :

$$p_\theta(x) = \frac{\theta^x}{(1+\theta)^{x+1}}, \quad x \in I\!N \text{ et } \theta \geq 0.$$

Nous avons $p_\theta^n(0) = (1+\theta)^{-n}$, $\frac{\partial}{\partial\theta}p_\theta^n(0) = -n(1+\theta)^{-n-1}$, $\frac{\partial^2}{\partial\theta^2}p_\theta^n(0) = n(n+1)(1+\theta)^{-n-2}$. Les conditions (i) à (v) sont toutes vérifiées. Nous utilisons alors le théorème 3.3.2 pour obtenir nos résultats de minimaxité. Ces résultats sont donnés par le corollaire suivant.

Corollaire 3.3.5. *Pour une distribution Géométrique, et pour une perte quadratique, l'estimateur de Bayes δ_m est minimax et π_m est la loi a priori la moins favorable si $m \leq m_0(n) = m_1(n) \wedge m_2(n)$ où $m_2(n) = 1/(n+1)$ et $m_1(n)$ est solution en m de l'équation $mn(1+m)^n + 2(mn-1)(1+m)^{n/2} - 2 = 0$.*

DÉMONSTRATION. Nous avons $z'(\theta) = n(n+1)(1+\theta)^{-n-3}[1-(n+1)\theta]$, donc $m_2(n) = 1/(n+1)$. Par le théorème 3.3.2, l'estimateur de Bayes δ_m est minimax pour $m \leq m_0(n) = m_1(n) \wedge m_2(n)$. Ce qui achève la démonstration. □

Remarque 3.3.4. *Nous savons, pour tout $n \geq 1$, que $h(m) \leq 0$ si et seulement si $m \leq m_1(n)$. Mais, nous avons :*

$$h(m_2(n)) = \frac{1}{n+1}\frac{n(n+2)^n - 2(n+2)^{n/2}(n+1)^{n/2} - 2(n+1)^{n+1}}{(n+2)^n + 2(n+2)^{n/2}(n+1)^{n/2}}.$$

Les évaluations numériques de $h(m_2(n))$, nous montrent que ce dernier est négatif pour $1 \leq n \leq 11$ et est positif pour $12 \leq n \leq 1000$. D'après ces évaluations numériques, nous pouvons dire que

$$m_0(n) = \begin{cases} m_2(n) \text{ pour } 1 \leq n \leq 11 \\ m_1(n) \text{ pour } 12 \leq n \leq 1000. \end{cases}$$

En utilisant alors (3) théorème 3.3.2, pour $12 \leq n \leq 1000$, la loi a priori π_m est la moins favorable et l'estimateur de Bayes δ_m est minimax si et seulement si $m \leq m_1(n)$ avec $m_1(n)$ défini par le corollaire précédent. Lorsque $1 \leq n \leq 11$, la loi a priori π_m est la moins favorable et l'estimateur de Bayes δ_m est minimax si $m \leq 1/(n+1)$.

Nous proposons par le tableau suivant, $m_0(n)$ pour différentes valeurs de n. Nous observons que l'espace des paramètres doit être assez petit pour que la loi a priori la moins favorable soit concentrée sur sa frontière. Nous posons à présent, *pour* $n \geq 1$, $y_n = nm_0(n)$ *et* $\lim\limits_{n\to\infty} y_n = y_{**}$. *Nous pouvons alors vérifier que* $\lim\limits_{n\to\infty}(y_n(1+y_n/n)^n + 2(y_n-1)(1+y_n/n)^{n/2} - 2) = 0$. *Alors,* y_{**} *serait solution de l'équation en* y *de* $y\exp(y) + 2(y-1)\exp(y/2) - 2 = 0$. *La valeur de* y_{**} *est approximativement égale à* 0.912955. *Lorsque* n *est assez grand,* $m_0(n) \approx 0.912955/n$.

TAB. 3.2. Valeurs de $m_0(n)$ pour une distribution Géométrique avec la perte quadratique.

n	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
$m_0(n)$	*0.50*	*0.33*	*0.25*	*0.20*	*0.167*	*0.143*	*0.125*	*0.111*	*0.10*	*0.09*
n	*11*	*20*	*30*	*40*	*50*	*60*	*70*	*80*	*90*	*100*
$m_0(n)$	*0.084*	*0.046*	*0.031*	*.023*	*0.018*	*0.015*	*0.013*	*0.011*	*0.010*	*0.009*

Remarque 3.3.5. *Nous pouvons noter que la limite de* $y_n = nm_0(n)$ *est la même pour les distributions Bernoulli, géométrique et de Poisson (c'est-à-dire que* $y_* = y_{**} = y_0$*). En effet, pour ces trois distributions, nous avons* $\lim\limits_{n\to\infty} p^n_{m_0(n)}(0) = \exp(-y_0)$. *En prenant la limite des membres de l'équation (3.3.16), nous avons* $y\exp(y) + 2(y-1)\exp(y/2) - 2 = 0$ *avec* $y = \lim\limits_{n\longrightarrow\infty} y_n$.

Nous allons à présent étudier, par la prochaine sous-section, le cas d'une distribution de Poisson pour une fonction de perte LINEX.

3.3.1.2. *Perte LINEX et distribution de Poisson*

Dans cette sous section, nous traitons le cas d'une distribution de Poisson. Sa fonction de masse est donnée par la relation (3.3.18). Nous considérons la fonction de perte LINEX donnée par la relation (3.2.3). Wan, Zou et Lee (2000) ont étudié ce problème et ont proposé, pour tout $n \geq 1$, des conditions suffisantes pour que la loi a priori π_m soit la moins favorable et l'estimateur de Bayes δ_m soit minimax. Ici, nous proposons, par le théorème 3.3.3 et le théorème 3.3.4 respectivement :

(1) pour $\gamma < 0$ et $n \geq 1$, des conditions nécessaires et suffisantes pour que la loi a priori π_m soit la moins favorable et l'estimateur de Bayes δ_m soit minimax.

(2) pour $\gamma > 0$ et $n \geq 1$, des conditions suffisantes pour que la loi a priori π_m

soit la moins favorable et l'estimateur de Bayes δ_m soit minimax. Il faut noter que les conditions suffisantes que nous proposons sont meilleures que celles proposées par Wan, Zou et Lee (2000).

Nos conditions nécessaires et suffisantes comme nos conditions suffisantes, s'obtiennent de façon similaire au cas d'une distribution exponentielle sous la perte LINEX. Il faut noter que les résultats obtenus dans le cas exponentiel et le cas Poisson sont similaires.

Théorème 3.3.3. *Pour tout $\gamma \neq 0$, nous avons les résultats suivants :*

(1) La fonction de risque de δ_m est convexe en θ si et seulement si $A_\gamma(m) \geq 0$ avec $A_\gamma(m) = \gamma^2 + (n+\gamma)^2(\exp(\gamma y(m) - m) - 1) + \gamma(m - y(m))\exp(-\gamma m)$. De plus, lorsque $\gamma > 0$, la condition $A_\gamma(m) \geq 0$ est équivalente à $m \leq m_1(n, \gamma)$ où $m_1(n, \gamma)$ est solution de l'équation $A_\gamma(m) = 0$ en m.

(2) Lorsque $\gamma > 0$, l'estimateur de Bayes δ_m est minimax et la loi a priori π_m est la moins favorable si $m \leq m_1(n, \gamma)$.

Démonstration. (1) La fonction de risque de δ_m est donnée par :

$$\begin{aligned} R(\theta, \delta_m) &= (\exp(\gamma(m-\theta)) - \gamma(m-\theta) - 1) + \exp(-n\theta)[\exp(\gamma(y(m) - \theta) \\ &\qquad - \exp(\gamma(m-\theta)) + \gamma(m - y(m))]. \end{aligned}$$

Nous pouvons écrire la deuxième dérivée de la fonction de risque de l'estimateur de Bayes δ_m de la façon suivante :

$$\begin{aligned} \frac{\partial^2}{\partial\theta^2} R(\theta, \delta_m) &= \exp(\gamma(m-\theta))[\gamma^2 + \exp(-n\theta)[(n+\gamma)^2(\exp(\gamma(y(m) - m)) - 1) \\ &\qquad + n^2\gamma(m - y(m))\exp(-\gamma(m-\theta))]], \qquad (3.3.18) \\ &= \exp(\gamma(m-\theta))A_{\gamma,m}(\theta), \end{aligned}$$

avec $A_{\gamma,m}(\theta) = \gamma^2 + \exp(-n\theta)B_{\gamma,m}(\theta)$,
où $B_{\gamma,m}(\theta) = (n+\gamma)^2(\exp(\gamma(y(m) - m)) - 1) + n^2\gamma(m - y(m))\exp(\gamma(\theta - m))$.
Mais, pour $\gamma \neq 0$, $B_{\gamma,m}(\theta)$ est croissante en θ. Nous avons alors $B_{\gamma,m}(\theta) \geq B_{\gamma,m}(0)$, pour tout $\theta \in \Theta(m)$. Ceci nous donne $A_{\gamma,m}(\theta) \geq \gamma^2 + \exp(-n\theta)B_{\gamma,m}(0)$, pour tout $\theta \in \Theta(m)$. Nous posons $A_\gamma(m) = A_{\gamma,m}(0)$. Nous distinguons deux cas.

(a) Si $B_{\gamma,m}(0)$ est positif alors $A_{\gamma,m}(\theta)$ est positif, pour tout $\theta \in \Theta(m)$.

(b) Si $B_{\gamma,m}(0)$ est négatif alors $A_{\gamma,m}(\theta) \geq A_{\gamma,m}(0)$, pour tout $\theta \in \Theta(m)$.

Nous avons alors $A_{\gamma,m}(\theta) \geq 0$ pour tout $\theta \in \Theta(m)$ si et seulement si $A_{\gamma,m}(0) \geq 0$. En outre, lorsque $\gamma > 0$, la fonction $A_\gamma(m)$ est décroissante en m avec $A_\gamma(0) = \gamma^2 \geq 0$ et $\lim_{m\to\infty} A_\gamma(m) < 0$. Ceci nous donne le résultat.

(2) C'est une conséquence directe du (1) et du théorème 3.3.8 (appendice). □

Pour obtenir nos conditions nécessaires et suffisantes, nous avons besoin du prochain lemme.

Lemme 3.3.3. *Sous la fonction de perte LINEX et pour $\gamma < 0$ et $n \geq 1$, nous avons le résultat suivant :*

$$\frac{\partial^3}{\partial\theta^3} R(\theta, \delta_\alpha) \geq 0, \quad \forall\, \theta \in \Theta(m) \text{ et } m \leq m^*(n,\gamma),$$

où $m^*(\gamma)$ est donné par le corollaire 3.3.2.

Démonstration. Par la relation (3.3.18), la dérivée troisième du risque de δ_m est donnée par :

$$\begin{aligned}\frac{\partial^3}{\partial\theta^3} R(\theta, \delta_m) &= -\exp(\gamma(m-\theta))[\gamma^3 + \exp(-n\theta)[-(n+\gamma)^3(\exp(\gamma(y(m)-m)-1)\\ &\quad -n^3\gamma(m-y(m))\exp(\gamma(m-\theta))]]\\ &= \exp(\gamma(m-\theta))[-\gamma^3 + \exp(-n\theta)H_\gamma(\theta,m)],\end{aligned}$$

où $H_\gamma(\theta,m) = -(n+\gamma)^3[\exp(\gamma y(m)-m)-1] - \gamma n^3(m-y(m)\exp(-\gamma(m-\theta)$.

(1) Si $\gamma \leq -n$, le résultat est immédiat.

(2) Nous supposons à présent que $-n < \gamma < 0$. La fonction $H_\gamma(\theta,m)$ est décroissante en θ. Nous avons alors :

$$\begin{aligned}H_\gamma(\theta,m) &\geq H_\gamma(m,m)\\ &= -(n+\gamma)^3[\exp(\gamma y(m)-m)-1] - \gamma n^3(m-y(m)),\\ &\geq -(n+\gamma)[\exp(\gamma y(m)-m)-1] - \gamma n(m-y(m)).\end{aligned}$$

En utilisant alors la condition nécessaire (3.3.17), nous avons $H_\gamma(\theta,m) \geq 0$, pour tout $\theta \in \Theta(m)$ et $m \leq m^*(n,\gamma)$. Ce qui achève la démonstration. □

Théorème 3.3.4. *Pour $\gamma < 0$, la loi a priori π_m est la moins favorable et l'estimateur de Bayes δ_m est minimax si et seulement si $m \in \{m : G_{\gamma,n}(m) \leq 0\} \neq \emptyset$, avec $G_{\gamma,n}(m) = \gamma + n\exp(\gamma m) - (n+\gamma)\exp(\gamma y(m)) - n\gamma(m - y(m))$.*

DÉMONSTRATION. Par le lemme précédent, nous avons la dérivée troisième de la fonction de risque positive pour tout $m \leq m^*(n,\gamma)$ et $\gamma < 0$. La condition (3.3.11) devient alors une condition nécessaire et suffisante pour que la fonction de risque atteigne son maximum en 0 et m. Mais, cette condition est équivalente à $G_{\gamma,n}(m) \leq 0$. Puisque $\frac{\partial}{\partial m}G_{\gamma,n}(m)\big|_{m=0} < 0$ et $G_{\gamma,n}(0) = 0$ alors l'ensemble $\{m : G_{\gamma,n}(m) \geq 0\}$ n'est pas vide. En utilisant alors le théorème 3.3.8 (appendice), nous avons le résultat. $\square$

Remarque 3.3.6. *Il est intéressant de remarquer que, pour tout $n \geq 1$ et $\gamma < 0$,*

$$\{m : A_\gamma(m) \geq 0\} \subseteq \{m : G_{\gamma,n}(m) \leq 0\}.$$

Conjecture 3.3.1. (1) *Pour tout $\gamma \neq 0$, la condition nécessaire (3.3.11) se réalise si et seulement si $m \leq m_0(n,\gamma)$ où $m_0(n,\gamma)$ est solution de l'équation $G_{\gamma,n}(m) = 0$ avec $G_{\gamma,n}(m) = \gamma + n\exp(\gamma m) - (n+\gamma)\exp(\gamma y(m)) - n\gamma(m - y(m))$.*
(2) *Pour $\gamma < 0$, la loi a priori π_m est la moins favorable et l'estimateur de Bayes δ_m minimax si et seulement si $m \leq m_0(n,\gamma)$.*
(3) *Pour tout $\gamma < 0$, $m_0(n,\gamma)$ est décroissant en n.*

Remarque 3.3.7. *La partie (2) de la conjecture est une conséquence de la partie (1) et du théorème 3.3.8 (apppendice).*
La partie (3) de la conjecture vient du fait que $G_{\gamma,n}(m)$ est croissant en n et de la partie (1). En effet si $n_1 \leq n_2$, alors $G_{\gamma,n_1}(m_0(n_2,\gamma)) \leq G_{\gamma,n_2}(m_0(n_2,\gamma)) = 0$. Par la partie (1) de la conjecture, nous avons $m_0(n_1,\gamma) \geq m_0(n_2,\gamma)$.

Nous vérifions, par l'exemple suivant, que la conjecture est vraie dans le cas particulier où $\gamma = -n$.

Exemple 3.3.6. *En utilisant la relation (3.3.7), pour $\gamma = -n$, $n \geq 1$ et $m \geq 0$, nous avons :*

$$y(m) = \frac{\exp(nm) - nm - 1}{n(\exp(nm) - 1)}.$$

La fonction $G_{-n,n}(m)$ est alors égale à :

$$\begin{aligned} G_{-n,n}(m) &= -n + n\exp(-nm) + n^2(m - y(m)), \\ &= \frac{n[-\exp(-2nm) + 3\exp(-nm) + nm - 2]}{1 - \exp(-nm)}. \end{aligned}$$

En se référant à l'exemple 3.2.4, nous avons $G_{-n,n} \leq 0$ *si et seulement si* $m \leq m_0(n,-n)$ *avec* $m_0(n,-n)$ *solution de l'équation en* m *de* $G_{-n,n}(m) = 0$. *Mais, si nous posons* $y = mn$ *alors l'équation* $G_{-n,n}(m) = 0$ *est équivalente à l'équation* $-\exp(-2y) + 3\exp(-y) + y - 2 = 0$ *et elle a pour solution* $y = 1.15139$. *Nous avons alors* $m_0(n,-n) = 1.15139/n$.

Pour finaliser cette section, nous présentons des évaluations numériques de la fonction de risque de δ_m, des fonctions $m_0(\gamma)$ et $m_1(\gamma)$ accompagnées de diverses commentaires.

(1) Lorsque $\gamma < 0$, nous savons que la condition nécessaire et suffisante pour la loi a priori soit concentrée sur $\{0, m\}$ est $G_{\gamma,n}(m) \geq 0$. Nous avons tenté en vain de démontrer, pour tout $n \geq 1$, que les conditions $G_{\gamma,n}(m) \leq 0$ et $m \leq m_0(n,\gamma)$ sont équivalentes avec $m_0(n,\gamma)$ solution de l'équation de $G_{\gamma,n}(m) = 0$. Mais, nous avons procédé à l'évaluation numérique de l'ensemble $\{m : G_{\gamma,n}(m) \geq 0\}$ et nous constatons que cet ensemble coïncide avec l'intervalle $[0, m_0(n,\gamma)]$. Nous savons aussi que la condition $A_\gamma(m) \geq 0$ est une condition nécessaire et suffisante pour laquelle le risque de δ_m est convexe, donc une condition suffisante pour que la loi a priori la moins favorable soit concentrée sur $\{0, m\}$. Nous avons évalué numériquement l'ensemble $\{m : A_\gamma(m) \geq 0\}$ et nous constatons que cet ensemble coïncide avec l'intervalle $[0, m_1(n,\gamma)]$ où $m_1(n,\gamma)$ est solution de l'équation $A_\gamma(m) = 0$, représenté à la figure 3.8.

(2) Lorsque $\gamma > 0$, nous savons que la condition $m \leq m_1(n,\gamma)$ est une condition nécessaire et suffisante pour laquelle le risque de δ_m est convexe en θ, donc une condition suffisante pour que la loi a priori la moins favorable soit concentrée sur $\{0, m\}$. Mais, nous avons aussi évalué numériquement les valeurs de $m_0(n,\gamma)$ pour lesquelles la loi a priori la moins favorable est concentrée sur $\{0, m\}$. Ces évaluations numériques sont basées sur l'algorithme décrit dans le cas d'une distribution exponentielle sous la fonction de perte LINEX (section 3.2.2).

(3) Nous avons donné, pour $n = 1$, quelques représentations graphiques de la fonction de risque de δ_m en fonction de θ pour différentes valeurs de m et γ, dans des cas où δ_m est minimax et dans des cas où δ_m n'est pas minimax, par les figures 3.6 et 3.7 respectivement. À partir de la figure 3.8, où sont représentées en fonction de γ, pour $n = 1$, 3, 5, 10, les valeurs de $m_0(n, \gamma)$ et $m_1(n, \gamma)$, nous constatons que

- $m_0(n, \gamma)$ décroît en fonction de γ. C'est-à-dire que plus γ est grand, plus l'intervalle, pour lequel la loi a priori la moins favorable est concentrée sur $\{0, m\}$ est petit. Par exemple, pour $n = 1$, si $\gamma = -3$, cet intervalle est donné par $[0, 1.951]$ et lorsque $\gamma = 3$, cet intervalle est $[0, 0.570]$.
- $m_1(n, \gamma)$ décroît avec γ et lorsque γ est grand, il faut que m soit assez petit pour que la loi a priori la moins favorable soit concentrée sur $\{0, m\}$. De plus, il faut noter que $m_0(n, \gamma) \approx 0.9129$ lorsque $\gamma \approx 0$ ce qui rejoint le résultat de minimaxité pour la perte quadratique que nous avons à la sous-section précédente (section 3.3.1.1).
- $m_0(n, \gamma)$ et $m_1(n, \gamma)$ décroissent en n. Lorsque n est grand, la loi a priori la moins favorable est concentrée sur $\{0, m\}$, si m est assez petit.

FIG. 3.1. Représentations graphiques de quelques fonctions $G_\gamma(m)$
en m pour $\gamma = 0.2$, 1, 5 et 10
avec $G_\gamma(m) = \gamma + (1-\gamma)\exp(\gamma(y(m)-m)) - \exp(-\gamma m) - \gamma y(m)$,
pour la perte LINEX : $L(\theta, d) = \exp(\gamma(d-\theta)) - \gamma(d-\theta) - 1$,
et la fonction de densité : $f(x|\theta) = \exp(-(x-\theta))I_{[\theta,+\infty)}(x)$.
Lorsque $\gamma > 0$, nous avons $\{m : G_\gamma(m) \geq 0\} = \{m : \delta_m \text{ minimax }\}$, avec
δ_m est l'estimateur de Bayes par rapport à la loi a priori qui concentre sa masse
sur les points $\{0, m\}$ et δ_m vérifie $R(0, \delta_m) = R(m, \delta_m)$.

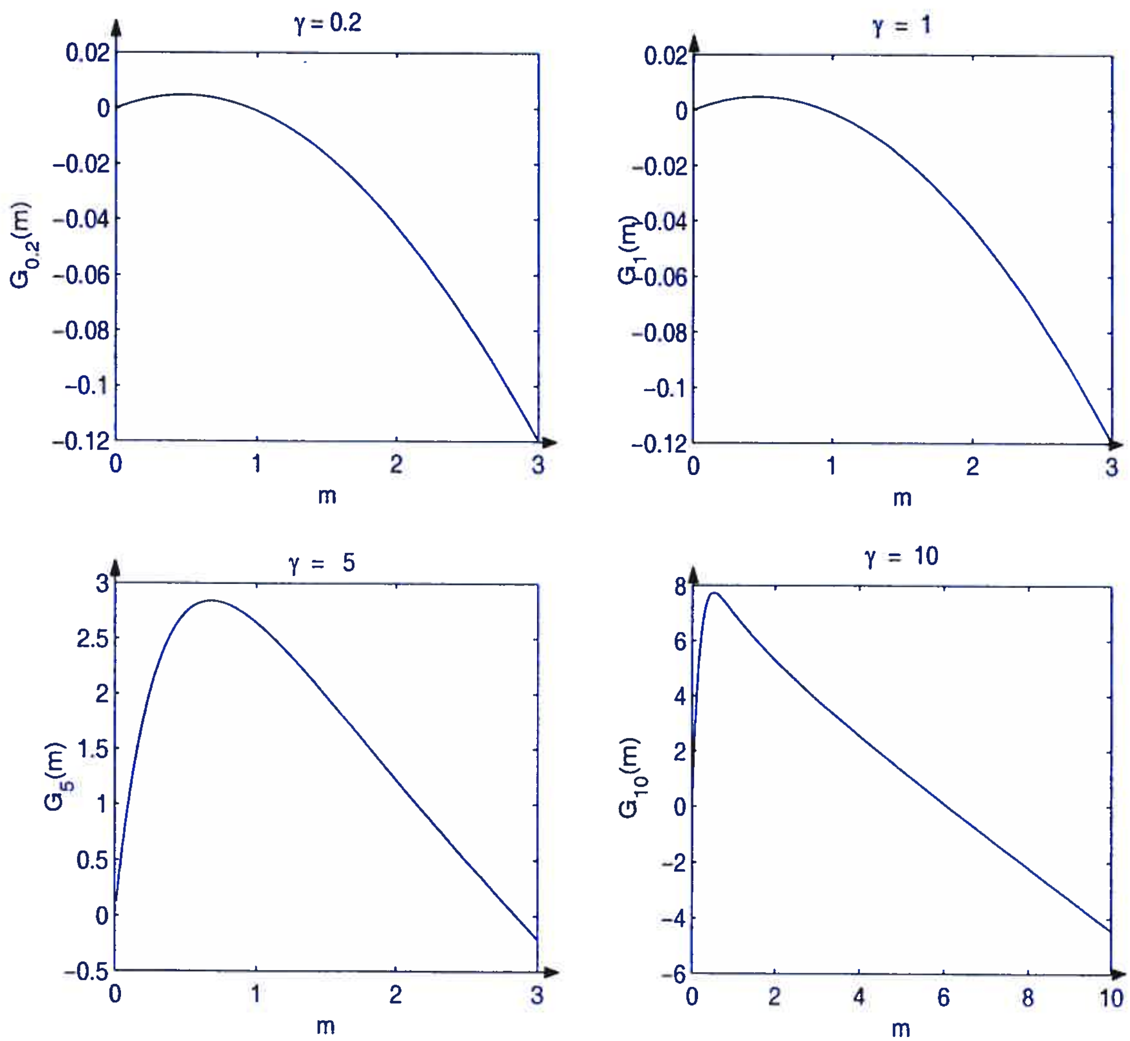


FIG. 3.2. Représentations graphiques de quelques fonctions de risque de δ_m en θ lorsque δ_m est minimax avec la perte LINEX : $L(\theta, d) = \exp(\gamma(d-\theta)) - \gamma(d-\theta) - 1$, et la fonction de densité : $f(x|\theta) = \exp(-(x-\theta))I_{[\theta,+\infty)}(x)$, δ_m est l'estimateur de Bayes par rapport à la loi a priori qui concentre sa masse sur les points $\{0, m\}$ et δ_m vérifie $R(0, \delta_m) = R(m, \delta_m)$.

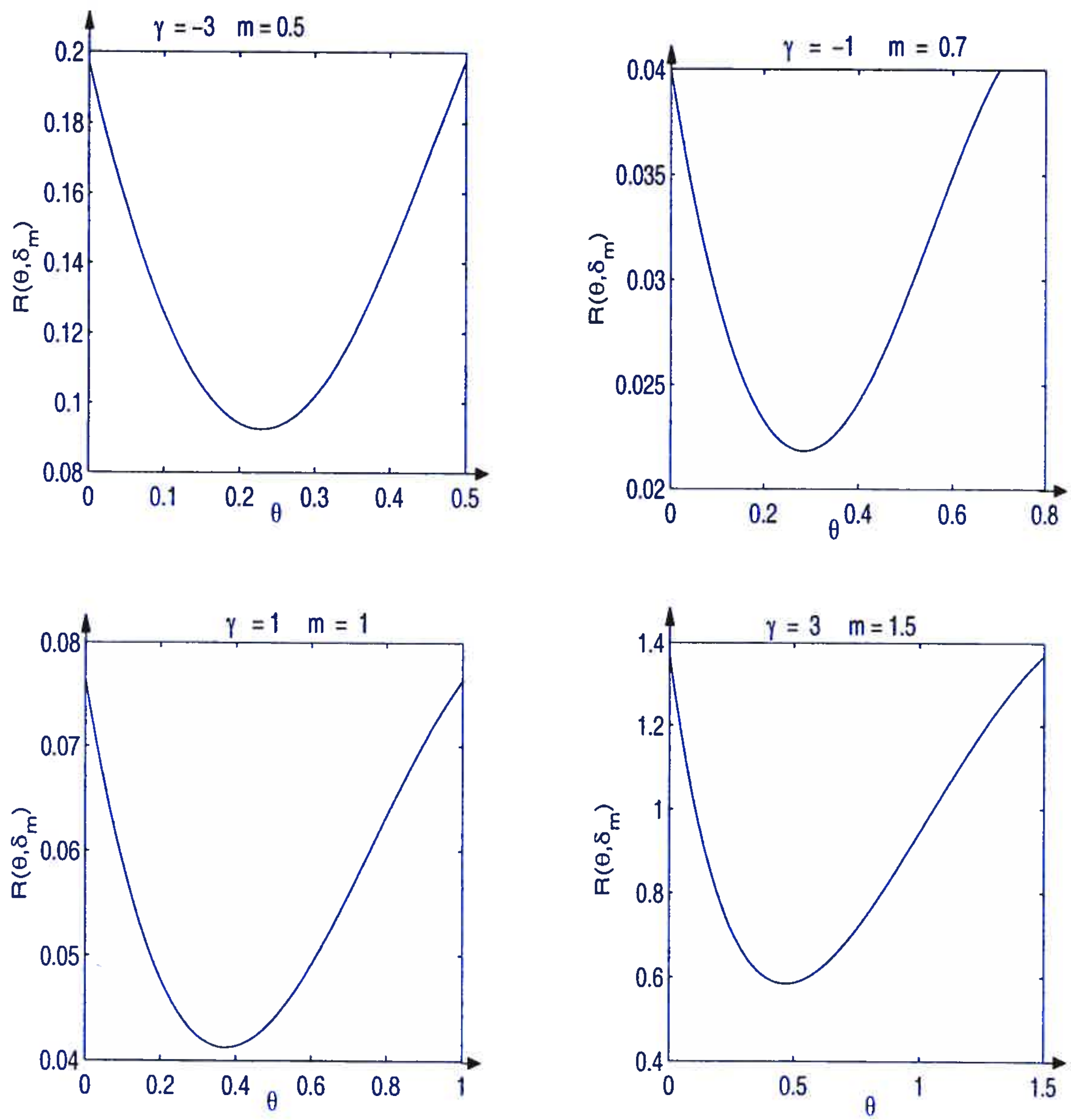


FIG. 3.3. Représentations graphiques de quelques fonctions de risque de δ_m en θ lorsque δ_m n'est pas minimax
avec la perte LINEX : $L(\theta, d) = \exp(\gamma(d-\theta)) - \gamma(d-\theta) - 1$,
et la fonction de densité : $f(x|\theta) = \exp(-(x-\theta))I_{[\theta,+\infty)}(x)$,
δ_m est l'estimateur de Bayes par rapport à la loi a priori qui concentre sa masse sur les points $\{0, m\}$ et δ_m vérifie $R(0, \delta_m) = R(m, \delta_m)$.

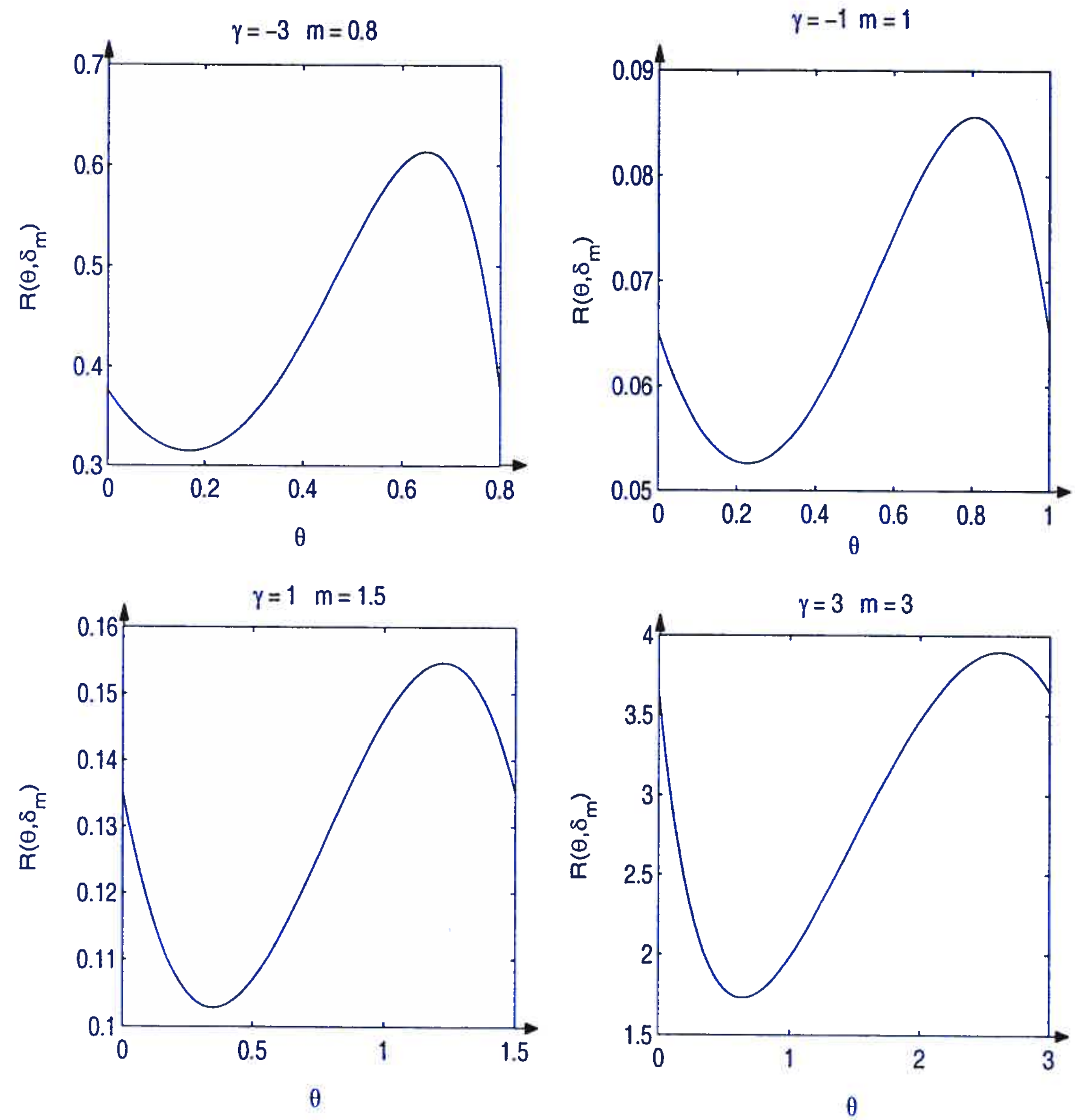


FIG. 3.4. Représentations graphiques de $m_0(\gamma)$ et $m_1(\gamma)$ en fonction de γ

avec la perte LINEX : $L(\theta, d) = \exp(\gamma(d-\theta)) - \gamma(d-\theta) - 1$

et la densité : $f(x|\theta) = \exp(-(x-\theta))I_{[\theta,+\infty)}(x)$.

Lorsque $\gamma > 0$: **(1)** nous avons $\{m : \delta_m \text{ minimax}\} = \{m : G_\gamma(m) \geq 0\}$ et $\{m : R(\theta, \delta_m) \text{ convexe en } \theta\} = \{m : A_\gamma(m) \geq 0\}$. **(2)**Par les évaluations numériques, nous avons $\{m : G_\gamma(m) \geq 0\} = \{m : m \leq m_0(\gamma)\}$ et $\{m : A_\gamma(m) \geq 0\} = \{m : m \leq m_1(\gamma)\}$ avec $m_0(\gamma)$ solution de l'équation $G_\gamma(m) = 0$ et $m_1(\gamma)$ solution de l'équation $A_\gamma(m) = 0$.

Lorsque $\gamma < 0$: **(3)** par les évaluations numériques, nous avons $\{m : \delta_m \text{ minimax }\} = \{m : m \leq m_0(\gamma)\}$. **(4)** Nous avons montré que $\{m : R(\theta, \delta_m) \text{ convexe en } \theta\} = \{m : A_\gamma(m) \geq 0\} = \{m : m \leq m_1(\gamma)\}$.

δ_m est l'estimateur de Bayes par rapport à la loi a priori qui concentre sa masse sur les points $\{0, m\}$ et δ_m vérifie $R(0, \delta_m) = R(m, \delta_m)$.

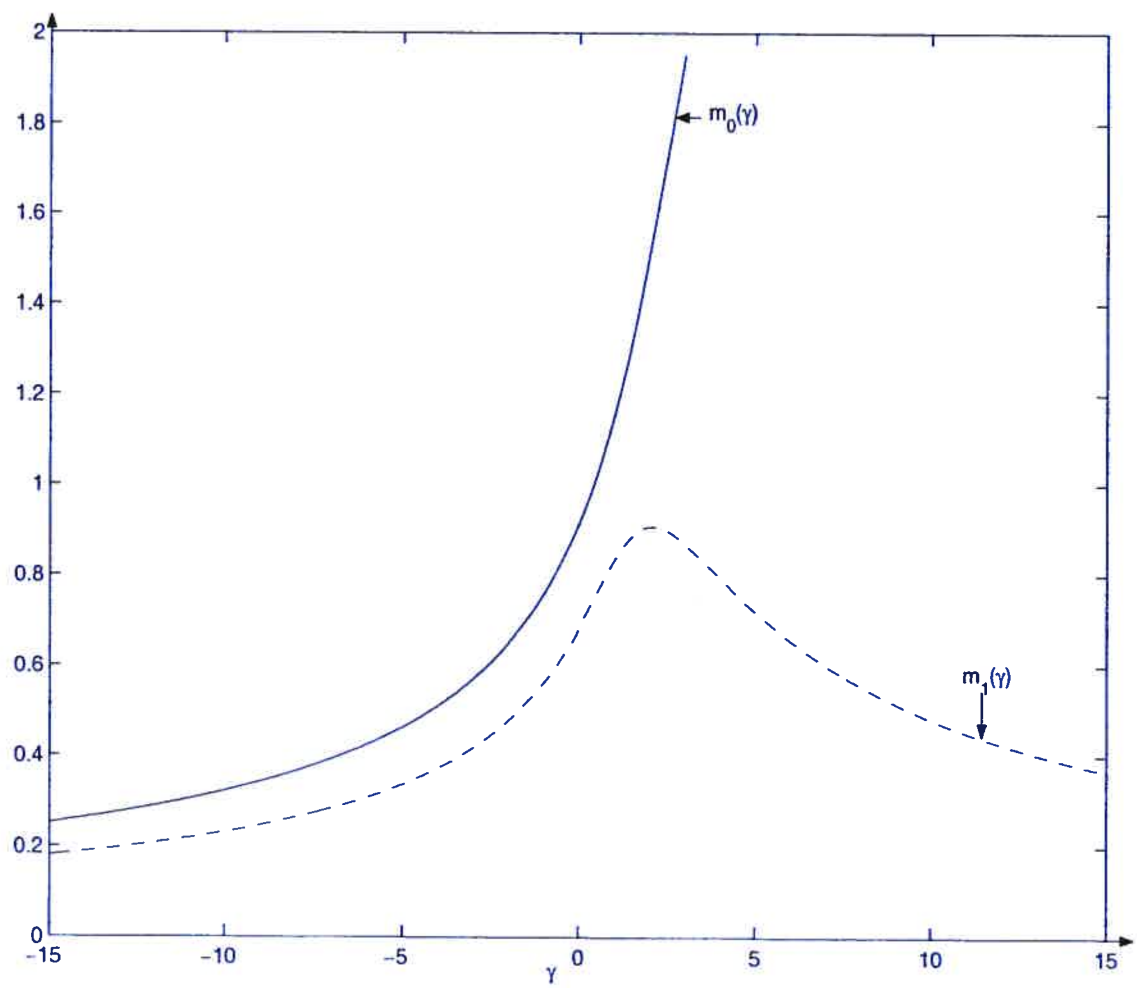


FIG. 3.5. Représentations graphiques de m_γ en fonction de γ avec la perte LINEX : $L(\theta, d) = \exp(\gamma(d-\theta)) - \gamma(d-\theta) - 1$ et la fonction de densité : $f(x|\theta) = c^{-1} I_{[0,c]}(x-\theta)$, $c \geq m$
Pour tout $\gamma \neq 0$, $\{m : \delta_m \text{ minimax}\} = \{m : m \leq m_\gamma\}$ avec m_γ solution de l'équation en m de $H_\gamma(m) = 0$ où
$H_\gamma(m) = |\gamma|\{-2(\exp(|\gamma|m) + \exp(-|\gamma|m) - 2) + |\gamma|m(1 - \exp(-|\gamma|m)) + |\gamma|^2 m(c-m)\}$
δ_m est l'estimateur de Bayes par rapport à la loi a priori qui concentre sa masse sur les points $\{0, m\}$ et δ_m vérifie $R(0, \delta_m) = R(m, \delta_m)$.

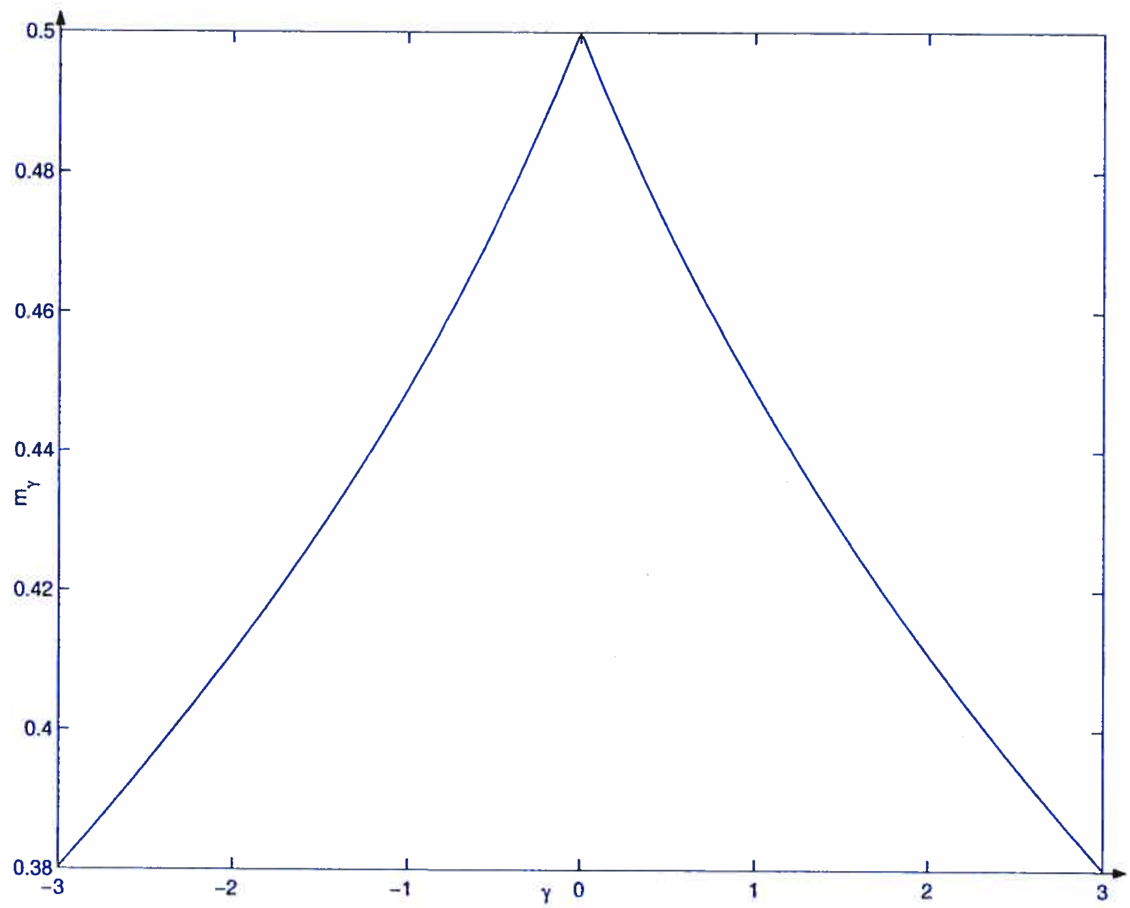


FIG. 3.6. Représentations graphiques de quelques fonctions de risque de δ_m en θ lorsque δ_m est minimax pour $n = 1$
avec la perte LINEX : $L(\theta, d) = \exp(\gamma(d-\theta)) - \gamma(d-\theta) - 1$,
et la fonction de masse : $p_\theta(x) = \exp(-\theta)\theta^x/x!$, $x = 0, 1, \ldots$
δ_m est l'estimateur de Bayes par rapport à la loi a priori qui concentre sa masse sur les points $\{0, m\}$ et δ_m vérifie $R(0, \delta_m) = R(m, \delta_m)$.

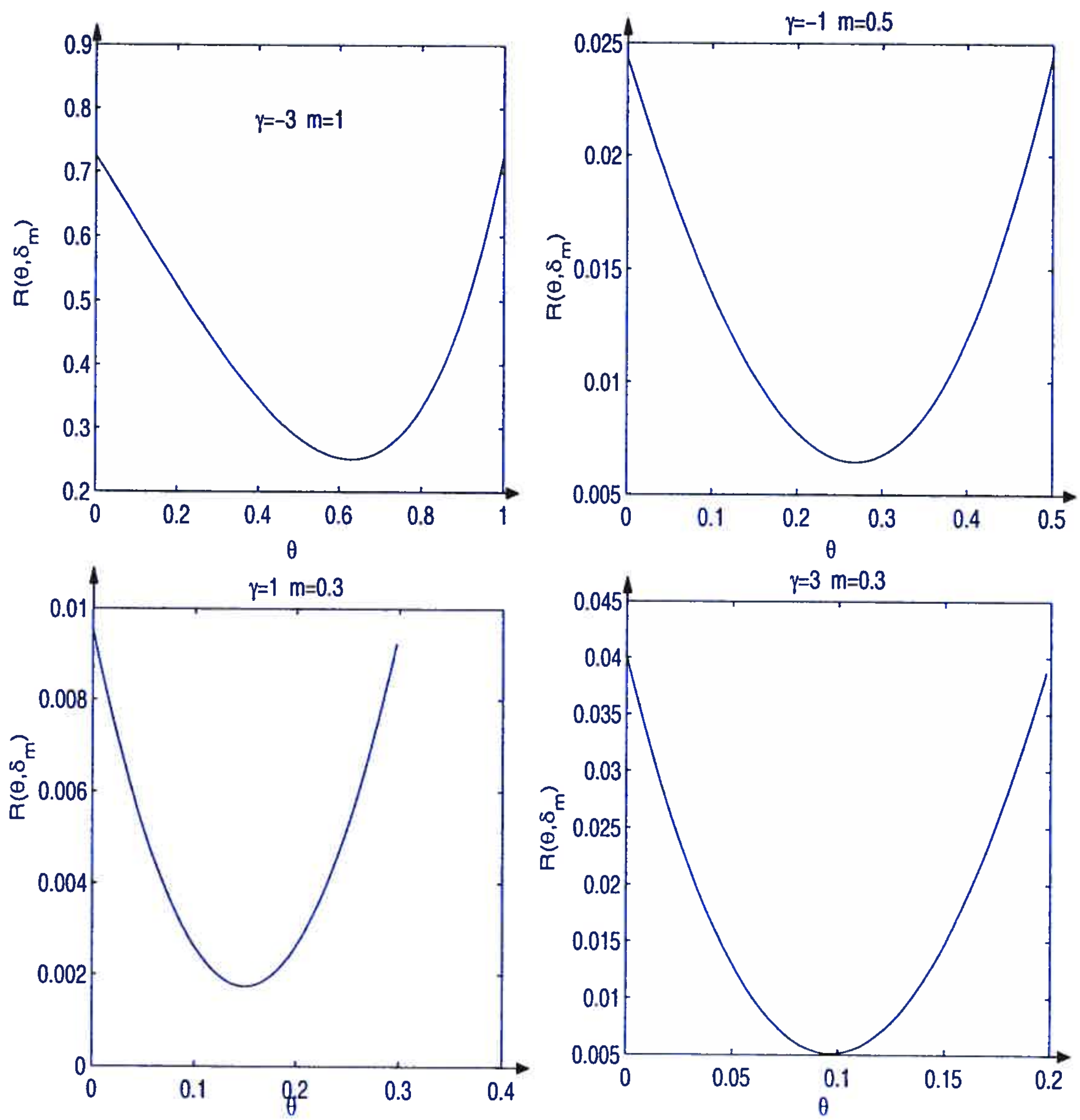


FIG. 3.7. Représentations graphiques de quelques fonctions de risque de δ_m en θ lorsque δ_m n'est pas minimax pour $n = 1$ avec la perte LINEX : $L(\theta, d) = \exp(\gamma(d - \theta)) - \gamma(d - \theta) - 1$, et la fonction de masse : $p_\theta(x) = \exp(-\theta)\theta^x/x!$, $x = 0, 1, \ldots$ δ_m est l'estimateur de Bayes par rapport à la loi a priori qui concentre sa masse sur les points $\{0, m\}$ et δ_m vérifie $R(0, \delta_m) = R(m, \delta_m)$.

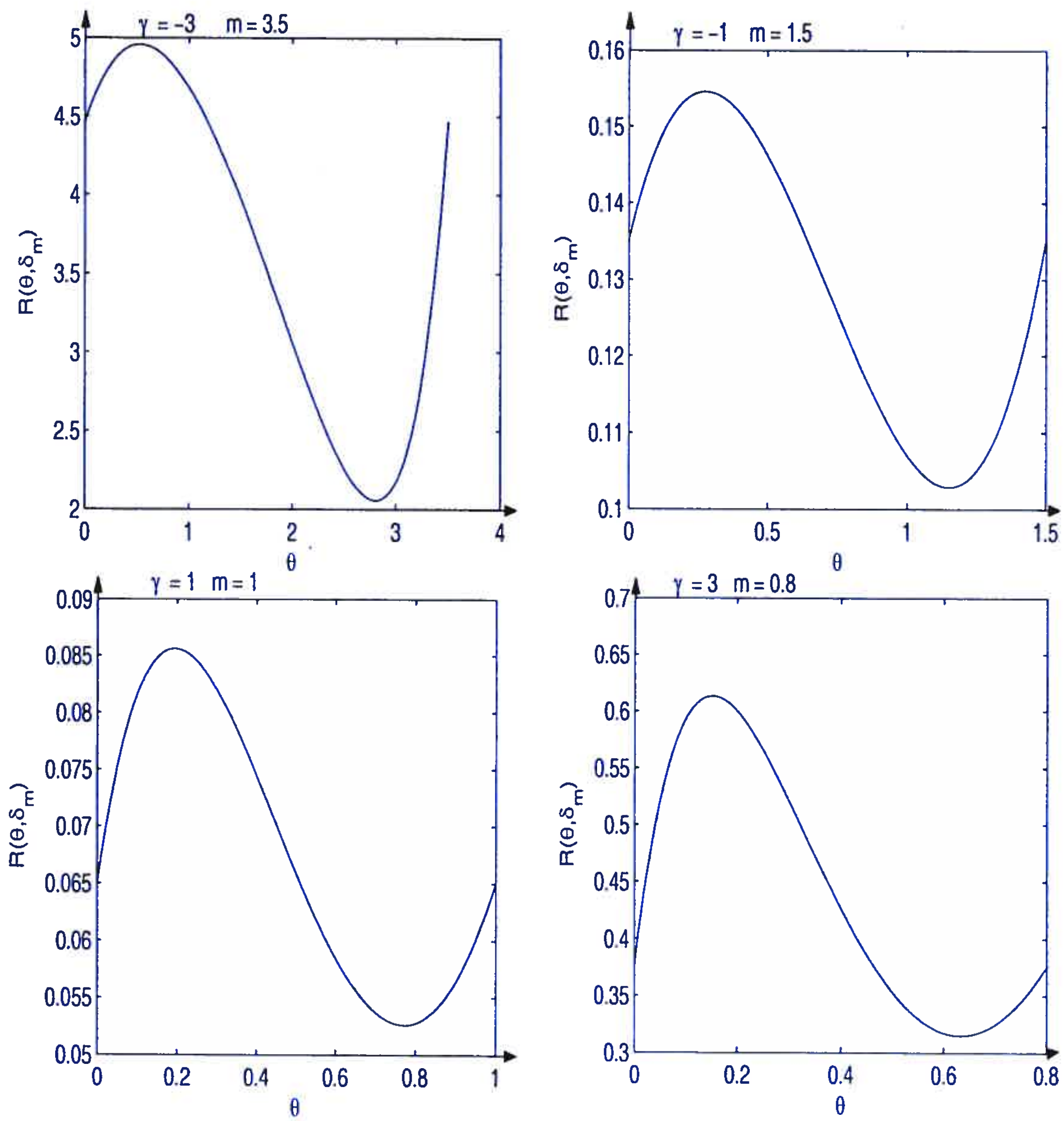


FIG. 3.8. Représentations graphiques de $m_0(\gamma, n)$ et $m_1(\gamma, n)$ en fonction de γ pour $n = 1, 3, 5, 10$

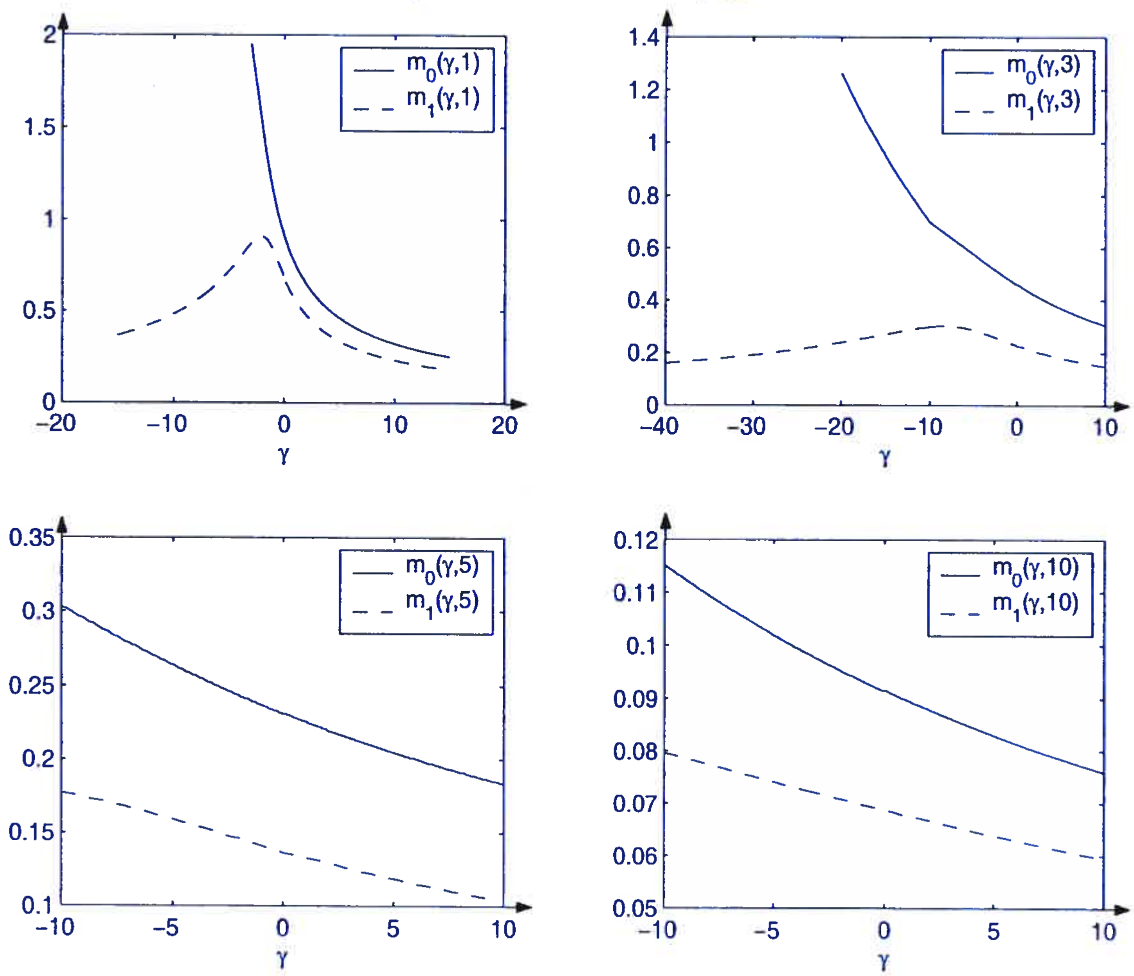
avec la perte LINEX : $L(\theta, d) = \exp(\gamma(d-\theta)) - \gamma(d-\theta) - 1$,

et la fonction de masse : $p_\theta(x) = \exp(-\theta)\theta^x/x!$, $x = 0, 1, \ldots$

Lorsque $\gamma < 0$: **(1)** nous avons $\{m : \delta_m \text{ minimax}\} = \{m : G_{\gamma,n}(m) \leq 0\}$ et $\{m : R(\theta, \delta_m) \text{ convexe en } \theta\} = \{m : A_\gamma(m) \geq 0\}$.**(2)**Par les évaluations numériques, nous avons $\{m : G_{\gamma,n}(m) \leq 0\} = \{m : m \leq m_0(\gamma, n)\}$ et $\{m : A_\gamma(m) \geq 0\} = \{m : m \leq m_1(\gamma, n)\}$ avec $m_0(\gamma, n)$ solution de l'équation $G_{\gamma,n}(m) = 0$ et $m_1(\gamma, n)$ solution de l'équation $A_\gamma(m) = 0$.

Lorsque $\gamma > 0$: **(3)** par les évaluations numériques, nous avons $\{m : \delta_m \text{ minimax }\} = \{m : m \leq m_0(\gamma)\}$. **(4)** Nous avons montré que $\{m : R(\theta, \delta_m) \text{ convexe en } \theta\} = \{m : A_\gamma(m) \geq 0\} = \{m : m \leq m_1(\gamma, n)\}$.

δ_m est l'estimateur de Bayes par rapport à la loi a priori qui concentre sa masse sur les points $\{0, m\}$ et δ_m vérifie $R(0, \delta_m) = R(m, \delta_m)$.



CONCLUSION

L'inadmissibilité de l'estimateur du maximum de vraisemblance pour le premier problème (premier chapitre) comme pour le deuxième problème (deuxième chapitre) est donné par les travaux de Charras et van Eeden (1991a). Dans le premier problème, si m est petit alors tout estimateur de Bayes par rapport à une loi a priori orthogonalement invariante domine l'estimateur du maximum de vraisemblance δ_{evm}. Lorsque m n'est pas petit, nous avons des séries de conditions sur le multiplicateur g d'un estimateur orthogonalement invariant pour que celui-ci domine δ_{evm}. Lorsque $m \leq c\sqrt{p}$ et p assez grand, nous montrons que l'estimateur de Bayes par rapport à la loi a priori uniforme sur $\Theta(m)$ domine δ_{evm}. Ces séries de conditions sont présentées dans le tableau 1.1. Aussi dans le deuxième problème, si m est assez petit alors tout estimateur δ symétrique autour de $1/2$ et qui rétrécit δ_{evm} autour de $1/2$ le domine. Lorsque m n'est pas petit, nous avons des séries de conditions pour lesquels δ_{evm} est dominé par un estimateur symétrique autour de $1/2$. Ces series de conditions sont données dans le tableau 2.1.

Pour les problèmes d'estimation minimax traités dans cette thèse, nous proposons parfois des conditions nécessaires et suffisantes, parfois des conditions suffisantes pour lesquelles la loi a priori la moins favorable soit concentrée sur la frontière de l'espace paramétrique. Ces conditions sont toujours de la forme $m \leq m_0$. Par exemple, dans le deuxième chapitre, nous avons obtenus des conditions suffisantes pour la minimaxité. Une voie de recherche pourrait être de trouver des conditions nécessaires et suffisantes sur m pour ce problème.

Dans le troisième chapitre, nous avons utilisé les fonctions de perte quadratique et LINEX pour proposer des solutions minimax. Une autre voie de recherche serait de considérer d'autres pertes asymétriques comme la perte
$L(\theta, d) = c_1(d-\theta)^2 I_{\{d\leq\theta\}} + c_2(d-\theta)^2 I_{\{d>\theta\}}$, pour tout c_1 et c_2. J'ai commencé à travailler sur ce problème et j'ai obtenu quelques résultats.

J'ai également travaillé le cas d'une distribution Binomiale et Géométrique avec l'utilisation d'une perte LINEX. Mais, j'ai obtenu quelques résultats similaires au cas d'une distribution de Poisson sous la perte LINEX. Ce qui me fait penser qu'il sera possible dans l'avenir de généraliser en utilisant une famille de distributions discrètes et une perte LINEX, pour trouver des solutions minimax. Il serait aussi intéressant de traiter le cas de la classe de distributions dite " Modified Power Series" pour une perte quadratique en essayant de vérifier les conditions (i) à (v) de la sous-section 3.3.1.1.

Il faut noter que Bader et Bischoff (2003) ont considéré le problème d'estimation minimax d'un paramètre θ lorsque ce dernier appartient à un intervalle $[\theta_0, \theta_0 + m]$, m et θ_0 connus. Ils traitent une grande classe de distributions continues avec des pertes strictement convexes. Ainsi, ils montrent l'existence de m_0 tel que pour $m \leq m_0$ la loi a priori la moins favorable est concentrée sur $\{\theta_0, \theta_0 + m\}$. Mais, ils ne donnent pas explicitement les valeurs de m_0.

Une dernière proposition de recherche serait une généralisation des résultats du deuxième chapitre avec l'utilisation d'une fonction de perte symétrique autre que les pertes traitées dans cette thèse chapitre.

APPENDICE

On considère $\mathcal{X}$ l'espace échantillonnal et $\mathcal{A}$ une tribu sur $\mathcal{X}$. Soit une famille de lois de probabilités $\mathcal{P}$ définies sur $\mathcal{A}$. Si le vecteur aléatoire X associé aux observations prend ses valeurs dans $\mathcal{X}$ et a pour loi de probabilité $\mathcal{L}(X) \in \mathcal{P}$, alors on dit que le triplet $(\mathcal{X}, \mathcal{A}, \mathcal{P})$ est un modèle statistique. Lorsque $\mathcal{P}$ s'écrit $\{\mathcal{P}_\theta, \theta \in \Theta\}$, Θ est dit l'espace paramétrique. On dira dans ce cas que le modèle statistique est paramétrisé et il est noté par $(\mathcal{X}, \mathcal{A}, \Theta)$. Une fois le modèle statistique construit, on cherche à établir une inférence sur le paramètre θ. On utilise l'observation x pour estimer la valeur du paramètre θ inconnue du statisticien. Le rôle de ce dernier consiste à tirer une conclusion concernant ce paramètre θ en utilisant toujours l'observation x. Cette conclusion est appelée décision en statistique. Elle est choisie parmi un ensemble de décisions noté $\mathcal{D}$. Cet ensemble est appelé l'espace des décisions. On munit cette espace d'une tribu $\mathcal{B}$ tel que $(\mathcal{D}, \mathcal{B})$ est un espace mesurable et que la tribu $\mathcal{B}$ soit assez grande pour contenir tous les singletons de $\mathcal{D}$.

Pour mesurer la précision d'une décision, on introduit une fonction appelée fonction de perte.

Définition 3.3.1. *On appelle fonction de perte (ou de coût espéré), la fonction L définie sur $\Theta \times \mathcal{D}$ à valeurs dans $\mathbb{R}^+$ telle que $\forall \theta \in \Theta$ la fonction suivante :*

$$L(\theta, \cdot) : D \longrightarrow \mathbb{R}^+,$$

est mesurable.

Cette notion est nécessaire au traitement des notions de minimaxité et admissibilité (définitions 3.3.4 et 3.3.5).

Définition 3.3.2. *On appelle règle de décision, toute fonction mesurable de $\mathcal{X}$ vers $\mathcal{D}$.*

Nous notons par Λ, l'ensemble des règles de décisions.
On désigne par $\{\mathcal{X}, \mathcal{A}, \Theta, \mathcal{D}, \mathcal{B}, L\}$ le modèle paramétrique décisionnel. On cherche à minimiser la fonction de coût par l'intermédiaire du coût moyen ou risque fréquentiste défini par :

$$\begin{aligned} R \ : \ \Theta \times \Lambda &\longrightarrow I\!R^{+} \cup \{\infty\} \\ (\theta, \delta) &\longrightarrow \int_{\mathcal{X}} L(\theta, \delta(x)) P_{\theta}(dx). \end{aligned}$$

Cette fonction est utilisée pour comparer deux règles de décisions. La meilleure règle est celle qui a le plus petit risque possible.

Définition 3.3.3. *Une règle δ_0 domine une autre règle δ_1 si $R(\cdot, \delta_0) \leq R(\cdot, \delta_1)$ et $R(\theta_0, \delta_0) < R(\theta_0, \delta_1)$ pour un certain paramètre $\theta_0 \in \Theta$.*

Définition 3.3.4. *Une règle δ_0 est inadmissible s'il existe une règle δ_1 telle que δ_1 domine δ_0. Dans le cas contraire, on dit que δ_0 est admissible.*

Ce critère de dominance induit un ordre partiel sur l'espace des décisions $\mathcal{D}$ en comparant les risques fréquentistes des estimateurs.

Le critère minimax présenté ci-dessous est en fait "une assurance contre le pire", au sens où il vise à minimiser les pertes dans le cas le plus défavorable.

Définition 3.3.5. *Une règle δ_0 est minimax si :*

$$\sup_{\theta \in \Theta} R(\theta, \delta_0) = \inf_{\delta \in \Lambda} \sup_{\theta \in \Theta} R(\theta, \delta).$$

C'est-à-dire qu'une règle est minimax si elle minimise le risque maximal.
Le risque de Bayes de la règle δ par rapport la probabilité a priori π est donné par la quantité suivante :

$$r(\pi, \delta) = \int_{\Theta} R(t, \delta) \pi(dt).$$

Il représente le risque moyen calculé à partir de la loi de probabilité a priori.

Définition 3.3.6. *Une règle δ_0 est une règle de Bayes par rapport à la loi a priori π si $r(\pi, \delta_0) = \inf_{\delta \in \Lambda} r(\pi, \delta)$.*

Il existe plusieurs relations entre la minimaxité et les règles de Bayes. La loi a priori la moins favorable (definition 3.3.7) sert principalement de lien entre la minimaxité et les règles de Bayes.

Définition 3.3.7. *Une loi a priori π_0 est la moins favorable si :*

$$\inf_{\delta} r(\pi_0, \delta) = \sup_{\pi} \inf_{\delta} r(\pi, \delta).$$

C'est-à-dire qu'une loi a priori est la moins favorable si elle maximise le plus petit risque de Bayes.

Les résultats suivants donnent des liaisons qui existent entre la minimaxité, la loi a priori la moins favorable et les estimateurs de Bayes.

Considérons les conditions suivantes données par Kempthorne (1987) :

A : Pour toute loi a priori π sur l'espace paramétrique Θ, l'estimateur de Bayes δ_π par rapport à π est unique, pour tout $\theta \in \Theta$ presque partout.

B : Si $\{\pi_i, i \geq 1\}$ est une suite de lois a priori sur Θ qui converge faiblement vers une loi π, alors les fonctions de risque $R(\cdot, \delta_{\pi_i})$, $i \geq 1$ des estimateurs de Bayes δ_{π_i} par rapport à π_i, convergent uniformément vers la fonction de risque $R(\cdot, \delta_\pi)$ de l'estimateur de Bayes δ_π par rapport à π.

C : l'espace paramétrique Θ est un espace métrique compact.

D : Pour tout estimateur δ, la fonction de risque $R(\theta, \delta)$ de δ est une fonction continue supérieurement en θ.

E : Pour tout estimateur δ de θ, la fonction de risque $R(\theta, \delta)$ de δ est une fonction analytique de θ.

Théorème 3.3.5. *Kempthorne (1987)*

(1) Supposons qu'une loi a priori π_0 vérifie les conditions A et B. Si π_0 est la moins favorable et δ_{π_0} est la règle de Bayes pour π_0 alors

$$\sup_{\theta \in \Theta} \quad R(\theta, \delta_{\pi_0}) = \inf_{\delta} \quad r(\pi_0, \delta). \tag{3.3.19}$$

(2) Si une loi a priori π_0 est telle que la relation (3.3.19) est vraie alors elle est la moins favorable.

Théorème 3.3.6. *Kempthorne (1987)*
Sous les conditions A à E, la loi a priori la moins favorable π_0 existe.

Théorème 3.3.7. *Kempthorne (1987)*
Sous les conditions A à E, la loi a priori la moins favorable vérifie au moins une des propriétés suivantes :

(1) la fonction de risque de l'estimateur de Bayes δ_{π_0} par rapport à π_0 est constante sur Θ,

(2) le support de π_0 est discret et fini.

Corollaire 3.3.6. *Kempthorne (1987)*
Si π_0 est une loi a priori la moins favorable alors l'estimateur de Bayes par rapport à π_0 est minimax.

Comme le montre le prochain lemme, pour qu'un estimateur de Bayes δ_{π_0} associé à la loi a priori π_0 soit minimax, il est nécessaire qu'il soit "égalisateur", c'est-à-dire que son risque soit constant sur le support de π_0.

Lemme 3.3.4. *Soit δ_{π_0} un estimateur de Bayes par rapport à la loi a priori π_0 discret sur les points $\theta_1, \theta_2, ..., \theta_k$ avec les probabilités respectives $p_1, p_2, ..., p_k$ où $p_i > 0$, $\forall\ i = 1, 2, ..., k$. Pour que la relation (3.3.19) soit vraie, il est nécessaire que $R(\theta_1, \delta_{\pi_0}) = R(\theta_2, \delta_{\pi_0}) = \cdots = R(\theta_k, \delta_{\pi_0})$.*

Démonstration. Le risque de Bayes de δ_{π_0} est donné par :

$$r(\pi_0, \delta_{\pi_0}) = \sum_{i=1}^{k} p_i R(\theta_i, \delta_{\pi_0}).$$

Supposons à présent que le risque $R(\theta_i, \delta_{\pi_0})$ de δ_{π_0} ne soit pas constant en θ_i. Soit j tel que :

$$\max_{1\leq i\leq k} R(\theta_i, \delta_{\pi_0}) = R(\theta_j, \delta_{\pi_0}).$$

Nous avons alors les relations suivantes :

$$\begin{aligned}
\sup_{\theta} R(\theta, \delta_{\pi_0}) \geq R(\theta_j, \delta_{\pi_0}) &= \sum_{i=1}^{k} p_i R(\theta_j, \delta_{\pi_0}) \\
&> \sum_{i=1}^{k} p_i R(\theta_i, \delta_{\pi_0}) \\
&\geq \inf_{\delta} r(\pi_0, \delta).
\end{aligned}$$

Ce qui veut dire que la relation (3.3.19) est fausse. □

Nous terminons par le théorème récapitulatif suivant qui donne non seulement une condition suffisante pour la minimaxité d'un estimateur δ_{π_0}, mais aussi une condition nécessaire basée sur la partie (2) du théorème 3.3.5 et l'unicité des règles de Bayes δ_{π_0}.

Théorème 3.3.8. *Étant donnée une loi a priori π_0, la règle δ_{π_0} associée à π_0 est minimax si et seulement si la relation (3.3.19) est vraie.*

BIBLIOGRAPHIE

AMOS, D. É. (1974). Computation of modified Bessel functions and their ratios. *Math. Computat.* **28,** 239-251.

BADER, G. ET BISCHOFF, W. (2003). Old and new aspects of minimax estimation of a bounded parameter. Mathematical Statistics and Applications : Frestschrift for Constance van Eeden, University of British Columbia, Marc Moore, Sorana Froda and Christian Léger, Editions Institute of mathematical Statistics Lecture Notes - Monographe Series 42, Institute of Mathematical Statistics, Beachwood, OH, 15-30.

BERRY, C. (1989). Bayes minimax estimation of Bernouilli p in a restricted parameter space. *Communications in Statistics series A* **18,** 4607-4616.

BERRY, C. (1990). Minimax Estimation of a Bounded Normal Mean Vector. *Journal of Multi. Anal.* **35,** 130-139.

BERRY, C. (1993). Minimax estimation of a restricted exponential location parameter. *Statistics and Decisions* **11,** 307-316.

BICKEL, P. (1981). Minimax estimation of the mean of a normal distribution when the parameter space is restricted. *Ann. of Math. Statist.* **9**, 1301-1309.

BISCHOFF, W., FIERGER, W. ET WULFERT, S. (1995). Minimax-and-Γ-minimax estimation of a bounded normal mean under LINEX loss. *Statistics and Decisions* **13,** 287-298.

BROWN, L. (1976). Notes on Statistical Decision theory. Unpublished Lecture Note, Ithaca.

CASELLA, G. ET STRAWDERMAN, W. (1981). Estimating a bounded normal mean. *Annals of Statistics* **9**, 870-878.

CHARRAS, A. ET VAN EEDEN, C. (1991A). Bayes and admissibility properties of estimators in truncated parameter space. *Canadian Journal of Statist.* **19,** 121-134.

CHARRAS, A. ET VAN EEDEN, C. (1991B). Limits of Bayes estimators in convex, truncated parameter spaces. *Statist. Probab. Lett.* **11,** 479-483.

CHARRAS, A. ET VAN EEDEN, C. (1992). Bayes properties of estimators of location parameters in truncated parameter spaces. *Statistics and Decisions* **10,** 81-86.

CHARRAS, A. ET VAN EEDEN, C. (1994). Inadmissibility for squared error loss when the parameter to be estimated is restricted to the interval $[a, \infty)$. *Statistics and Decisions* **12,** 257-266.

DASGUPTA, A. (1985). Bayes minimax estimation in multiparameter families when the parameter space is restricted to a bounded convex set. *Sankhya, Ser. A* **47,** 326-332.

EICHENAUER, J. (1986). Least favourable two point priors in estimating the bounded location parameter of a noncentral exponential distribution. *Statistics and Decisions* **4,** 389-392.

EICHENAUER, J. ET FIEGER, W. (1992). Minimax estimation under convex loss when the parameter interval is bounded. *Metrika* **39,** 27-43.

FERGUSON, T. (1967). *Mathematical Statistics : a decision-Theoretic approach.* Academic-Press, NY.

FUNO, E. (1991). Inadmissibility results of the MLE fort he multinomial problem whenthe parameter space is restricted or truncated. *Communications in Statistics : Theory and Methods* **20,** 2863-2880.

GHOSH, M. N. (1964). Uniform approximation of minimax estimates. *Ann. of Math. Statist.* **35,** 1031-1047.

JOHNSON, B. MCK. (1971). On the dmissible estimators for certain fixed sample binomial problems. *Ann. of Math. Statist.* **42,** 1579-1587.

JOHNSON, N.L. ; KOTZ, S. ET KEMP, A.W. (1992) *Univariate discrete distributions,* second edition, Wiley, New York.

JOHNSTONE, I. ET MACGIBBON B. (1992). Minimax estimation of a contrained Poisson vector. *Annals of Statistics* **20,** 807-831.

KARLIN, S. (1957). Pòlya type distributions, II. *Ann. of Math. Statist.* **28,** 281-309.

KEMPTHORNE (1987). Numerical specification of discrete least fovorable prior distributions. *SIAM Journal of Scientific Computing* **8,** 171-184.

KEMPTHORNE (1988). Dominating inadmissible procedures using compromise decision theory. *Statist. Decision Theory and Related Topic IV,* Springer-Verlag, New York, 381-396.

KHATTREE, R. (1992). Estimation of guarantee time and mean life after warranty for two parameter exponential failure model. *Austral. Journal Statist.* **34,** 207-215.

LEHMANN, E.L. ET CASELLA, G. (1998) *Theory of Point Estimation,* Springer-Verlag, New York.

MARCHAND, É. ET PARSIAN, A. (2002). On a Boundary Least Favourable Prior For Discrete Families and General Convex Loss. *Personnal communication.*

MARCHAND, É. ET MACGIBBON, B. (2000). Minimax estimation of a constrained Binomial proportion. *Statistics and Decisions* **18,** 129-167.

MARCHAND É. ET PERRON F. (2001). Improving on the MLE of a Bounded Normal Mean. *The Annals of Statist.* **29**, 1078-1093.

MARCHAND É. ET PERRON F. (2002). On the minimax estimator of a bounded normal mean. *Statist. and Probab. Lett.* **58**, 327-333.

MOORS, J.J.A. (1981). Inadmissibility of linearly invariant estimators in truncated parameter spaces. *Journal of the American Statisti. Assoc.* **76,** 910-915

MOORS, J.J.A. (1985). *Estimation in truncated parameter spaces.* Ph.D thesis, Tilburg University.

PARSIAN, A. ET KIRMANI, S. N. U. A. (2002). Estimation under LINEX Loss Function. *Handbook of applied econometrics and statistical inference,* 53-76, Statist. Textbooks Monogr., Dekker, New York.

PERRON, F. (1987). *Application de l'invariance en théorie de la décision.* Thèse de doctorat, Département de mathématiues et de statistique, Université de Montréal.

PERRON, F. (2003). Improving on the MLE of p for a Binomial(n, p) when p is around 1/2. Mathematical Statistics and Applications : Frestschrift for Constance van Eeden, University of British Columbia, Marc Moore, Sorana Froda and Christian Léger, Editions Institute of mathematical Statistics Lecture Notes - Monographe Series 42, Institute of Mathematical Statistics, Beachwood, OH, 45-64.

ROBERT, C. (1990). Modified bessel functions and their applications in probability and statistics. *Statist. and Probab. Lett.* **9,** 155-161.

ROBERT, C., BOCK M.E. ET CASELLA, G. (1990). Bayes estimators associated with uniform distributions on spheres (II) : the hierarchiecal Bayes approach. *Tech. Report BU-1002-M,* Cornell University.

SACKS, J. (1963). Generalized Bayes solutions in estimation problems. *Annals of Mathematical Statistics* **34,** 751-768.

VARIAN, H.R. (1975). *A Bayesian approach to real estate assessment.* In Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage, eds. Stephen E. Fienberg and Arnold Zellner, Amsterdam : North-holland, pp. 195-208.

WALD, A. (1950). *Statistical Decision Functions.* Wiley, New York.

WAN, A. T. K., ZOU, G. ET LEE, A. H. (2000). Minimax and Γ-minimax estimation for the Poisson distribution uneder the LINEX loss when the parameter space is restricted. *Statist. Probab. Lett.* **50,** 23-32.

WARNER, S. L. (1965). Randomized Response : A survey Technique for Eliminating Evasive Answer Bias, *Journal of the American Statistical Association* **60,** 63-69.

ZELLNER, A. (1986). Bayesian estimation and prediction using asymmetric loss functions. *Journal Amer. Statist. Assoc.* **81,** 446-451.

WATSON, G. S. (1983). *Statistical on spheres.* Wiley, New York.